Donders Graduate School for Cognitive Neuroscience

Master of Science Programme

MSc Thesis

Replicability and Uniqueness of Tremor Characteristics in Parkinson's Disease

Kelsey van Dun[1]

[1]*Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour, the Netherlands*

Supervised by:

1. Rick Helmich[1]
2. Eva Klimars[1]

Second reader:

1. Ian Cameron[1]

Abstract

Parkinson's disease (PD) is a heterogeneous disease characterized by bradykinesia, rigidity, and resting tremor. While bradykinesia and rigidity have a consistent response to mediation, resting tremor has not. Research into the causes of heterogeneous responses of tremor to medication is necessary, as patients describe tremor as their second most bothersome symptoms, and clinicians are currently forced to follow a trial-and-error approach when treating this symptom. Therefore, this study aims at identifying replicable and unique tremor-characteristics that might be used to identify subpopulations of PD to predict the most suitable treatment. We measured patients with PD on two occasions. We assessed replicability of the clinical severity of the tremor, frequency and amplitude in rest and during cognitive-coactivation, the spatial correlation within the tremor-amplitude related network, and the stability of the peak activation location within this network. We observed high replicability of clinical- and accelerometry characteristics. Spatial correlations and Euclidean distances were not replicable between days. We conclude that clinical- and accelerometry parameters might be suitable for identification of subgroups of patients, but our fMRI-derived parameters seem uninformative for clinical practice.

Keywords: Parkinson's disease, replicability, cerebello-thalamo-cortical-network, biomarkers, resting tremor

Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by prominent motor symptoms such as rigidity, bradykinesia, and resting tremor. Tremor is one of the cardinal symptoms of PD and is reported in almost seventy-five percent of patients. It is defined as an involuntary, rhythmic, and oscillatory movement of a body part, most often the limbs (Deuschl *et al.,* 1989). PD tremor often starts unilateral and can stay asymmetrical through disease duration. The typical PD tremor occurs at rest, meaning that the affected body part is not voluntarily activated and is completely supported against gravity. Although multiple other types of tremor might occur in PD, this paper will primarily focus on resting tremor.

PD is hallmarked by degeneration of dopaminergic cells in the substantia nigra pars compacta (SNpc), an area in the midbrain (Fearnley & Lees, 1991). The motor symptoms in PD are thought to result from this degeneration, leading to less dopaminergic input to the striatum (Kish, Shannak & Hornykiewicz, 1988; Bezard, Dovero, & Prunier, 2001). PD is thus characterized by lower levels of striatal dopamine. However, dopaminergic loss in the striatum only correlates with two of three main clinical PD motor symptoms, bradykinesia and rigidity, but not resting tremor (Helmich, Janssen, Oyen, Bloem, & Toni, 2011). Conversely, tremor severity correlated specifically with dopamine depletion in the globus pallidus internus (GPi), globus pallidus externus (GPe), and the putamen. Helmich *et al.*, (2011) also observed that the amplitude of resting tremor is associated with another circuit involving the ventral intermediate nucleus of the thalamus, the motor cortex, and the cerebellum, also called the CBLM-VIM-MC-circuit or the cerebello-thalamo-cortical-circuit. Pallidal depletion has been suggested to underlie increased coupling between the basal ganglia and the CBLM-VIM-MC-circuit (Helmich *et al.*, 2011; Helmich, Hallet, Deuschl, Toni, & Bloem, 2012). Therefore, tremor is currently interpreted as the pathological

interaction between the basal ganglia and the cerebello-thalamo-cortical circuit, in which the basal ganglia initiate resting tremor and the CBLM-VIM-MC-circuit propagates it.

Further, resting tremor in particular shows a highly variable and inconsistent response to available medications across patients. To date, most medication is dopaminergic-based, with levodopa as the most common prescribed one. Levodopa is considered the most effective treatment so far, but still only reduces tremor in a subgroup of patients**.** Dirkx *et al* (2017) investigated how dopamine influences the CBLM-VIM-MC-circuit and the basal ganglia. They observed that dopamine reduced activity in the GPi and VIM, but more importantly, directly increased the self-inhibition of the VIM, thereby reducing tremor. This effect was specifically found for resting tremor, and was thus not associated with bradykinesia and rigidity. Moreover, this effect was only present for dopamine-responsive tremor. Further research into the causes of heterogeneous responses of tremor to medication is necessary, as patients describe tremor as their second most bothersome symptoms, and clinicians are currently forced to follow a trial-and-error approach when treating this symptom.

All the above findings are based on group-level analyses, where between-subject differences are not considered. Although the role of the cerebello-thalamo-cortical circuit is proven to be highly consistent and replicable across independent cohorts of PD patients (Helmich *et al*., 2011; Helmich *et al*., 2012; Dirkx *et al*., 2017), it has not been investigated at an individual level. It therefore remains unclear to what extent inter-individual variability in the architecture of this circuit exists and how this might relate to differences in treatment response, e.g. help with predicting which treatment would be most effective. A requirement to use information about the architecture of the tremor circuit as a predictor of treatment success, is that parameters show sufficient interindividual differences and at the same time intraindividual stability over time.

There is increasing evidence that functional differences in brain networks between individuals exist. Finn *et al* (2015) examined both resting-state and task-related data of the Human Connectome Project (HCP) to investigate the predictive value of individual functional connectivity networks. Healthy participants were scanned on consecutive days. The authors show that it is possible to use functional connectivity parameters as a predictor to identify an individual from a large group, with accuracies between 96-97% when using resting-state data. However, the between-session interval was one day and the data was of very high quality thereby making it not directly generalizable. Nevertheless, a replication study with lower quality data was also able to identify individuals based on their functional connectivity profiles, albeit with a lower accuracy around 55-65% (Waller *et al.*, 2017). This might indicate that brain networks are unique to an individual and stable over time. However, this has not been tested specifically for the CBLM-VIM-MC network in PD patients.

Moreover, resting tremor is often quantified using clinical- and electrophysiological measurements, focusing mostly on tremor frequency and tremor power. Tremor frequency often ranges from 4-6 Hz in PD (Budzianowska & Honczarenko, 2009). It is known that tremor frequency is relatively stable over time when comparing groups (Hellwig *et al*, 2009). Results regarding stability of tremor amplitude are unclear, but at least indicate that it is highly context-dependent (Zach *et al*, 2015). However, the between-subject variability and long-term stability of these characteristics have also not been tested yet.

Therefore, this study aims at identifying the stability and uniqueness of tremor-related parameters on an individual level. To address this question, we move from group-level inference to individual data, and use precision neuroscience to work towards personalized medicine. We aim to identify tremor-related characteristics that might serve as an individual tremor *fingerprint*, based on the combination of individual neuroimaging, accelerometry- and clinical data. These individual characteristics might then help in identifying subpopulations of

patients, and in predicting treatment success and selecting patient-tailored treatments. We hypothesize that there are individual differences in all three modalities and that it is possible to establish an individual tremor *fingerprint* that is stable over time, which might function as a biomarker. We test this hypothesis by using concurrent accelerometry (ACC) with resting-state fMRI while focusing on the earlier established CBLM-VIM-MC network. We additionally extract tremor characteristics from behavioral accelerometry- and clinical measures.

Materials and Methods

Subjects and Inclusion

We recruited patients through their neurologist at the Neurology department of the Radboud University Medical Centre in Nijmegen. Most patients had participated in Parkinson's research previously and agreed to be subsequently approached for further research. We included 13 patients with a diagnosis of idiopathic Parkinson's disease according to the UK brain bank criteria with the presence of a clear resting tremor of at least one arm. Exclusion criteria were 1) the presence of other psychiatric or neurological symptoms, 2) presence of a severe head tremor or dyskinesia and 3) contraindications for (f)MRI scanning and transcranial alternating current stimulation (tACS). Patients had to be compatible for tACS, because as a part of this study patients also underwent tACS over the primary motor cortex at individual tremor frequency to assess whether tremor could be modulated by this non-invasive type of brain stimulation. However, this research question will not be addressed in this thesis.

Patients were measured on two separate days, with on average 78 days in between. On both days the exact same measurements were conducted, with the addition of two short tACS sessions on the second day only: a 25 minute tACS session (total stimulation duration 16 minutes) halfway the testing day and a 30 minute tACS-fMRI session (total stimulation

duration 20 minutes) after all other assessments. Measurements started in the morning and lasted approximately three and a half hours on day one and six hours on day two. Patients were measured in the practically defined off medication state, i.e. ≥ 12 h off short acting levodopa, ≥ 24 h off long acting levodopa and short acting dopamine agonists, and ≥ 48 h off extended release dopamine agonists. All participants gave written informed consent prior to their inclusion.

Cognitive Tests

We assessed cognitive functioning using the Montreal Cognitive Assessment (MOCA) and the Frontal Assessment Battery (FAB) only on day one as an indication of overall cognitive performance. Scores were not used as an exclusion criterion.

Clinical Tremor Measurements

Our first outcome measure is the clinical resting tremor score of the most affected arm as part of the Fahn-Tolosa-Marin Tremor Rating Scale Part A (Fahn *et al.,* 1993), which was used to evaluate tremor on both days. Tremor was rated by the same rater on both days. Additionally, the severity of motor symptoms was assessed on day one using the UPDRS part III. Videos were made to allow for a second opinion from a neurologist in case of uncertainty about ratings.

Tremor Activity Assessment

During the following assessment of tremor activity using accelerometry, we identified a number of individual tremor characteristics, such as the rest tremor frequency, rest tremor power (log transformed) and the change in both parameters during cognitive co-activation (mental arithmetic). On both days an identical procedure was followed. We quantified tremor movements using a three dimensional Acceleration Sensor (ACC; Brain Products GmbH, Germany; sensitivity: 420 mV/g) with standardized placement: centrally on the dorsum of the

most affected hand. This standardization allows us to compare outcome measures on different days to each other. The tremor of patients was assessed for two conditions while participants were sitting in a chair with their feet on the ground, their back supported, the lower arms lying on the armrests and their hands unsupported and unrestricted. We recorded tremor in rest and during cognitive co-activation (coco), which consisted of loud backwards counting in steps of three or seven as fast as possible, i.e. in the individual maximum speed while the examiner verbally encouraged rapid responding (social evaluation). The rest and coco condition were alternated and measured three times, each trial lasting one minute. The procedure was recorded on video and recorded using a BrainAmp ExG and BrainVision Recorder (Brain Products GmbH, Germany).

Processing of Tremor Accelerometry

Accelerometry data was preprocessed using Fieldtrip (Oostenveld et al, 2011). Data was detrended, demeaned, and bandpass filtered at [2-20 Hz]. Data was segmented in five seconds and a Fast Fourier Transform was applied to frequencies between 2-16 Hz, using a Hanning Taper of 2 seconds with a resolution of 0.5 Hz. All timepoints were averaged to extract the average powerspectrum. We then calculated the peak tremor frequency and the log-transformed amplitude at peak frequency for both the rest and coco condition. Log-transformed rest amplitude values were subtracted from the log transformed coco amplitude value to indicate sensitivity to cognitive stress.

Additionally, accelerometry measurements during fMRI were processed identically to the tremor-accelerometry data mentioned above with additional calculation of the time course of the tremor amplitude over time at individual tremor frequency. We also calculated the first derivative of the amplitude regressor. This resulted in two patient-specific regressors, one describing fluctuations in tremor amplitude, and one describing onset of tremor episodes.

After convolution with a haemodynamic response function, we used both regressors in a multiple-regression analysis.

Image Acquisition

In a third assessment we measured tremor-related activity using concurrent accelerometry-fMRI scanning. We obtained an fMRI scan with the same ACC setup as during the tremor measurement mentioned above. The scan on day one was identical to day two, but due to the additional tests performed on day two the measurement was on average about 1.5 hours later. Additionally, the second fMRI acquisition was conducted with MRI-compatible tACS electrodes (one above the motor cortex and one on the shoulder), which is not considered in this paper. This did not cause major artefacts in the data; sporadic distortions were limited to the skull. A 3T Siemens Magnetom MRI system with a 64-channel head coil was used to scan all 13 subjects. We conducted a resting-state fMRI scan where patients were instructed to lie still with their eyes open while focusing on a white fixation cross on a black screen. It was emphasized that patients had to stay awake and had to keep their most affected hand as unrestricted as possible. We used an interleaved multiband gradient-echo echo-planar imaging (GE-EPI) acquisition with an acceleration factor of 6 (TE = 34ms, TR=1s, 72 axial slices**,** voxel size 2mm x 2mm x 2mm, field of view **=** 210 mm, flip angle = 60 degrees, scanning time = 10 minutes; 600 images). Additionally, on day one we obtained an anatomical MPRAGE T1-weighted scan for registration of the functional data of both sessions. (TR = 2300 ms, TE = 3 ms, voxel size = 1mm x 1mm x 1mm, flip angle= 8 degrees, FOV =256 mm**,** scanning time = 10 minutes**)**.

Preprocessing of fMRI data

Preprocessing was done using FMRIB Software Library (FSL 6.0, Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012). Functional preprocessing was performed

using FMRI Expert Analysis Tool (FEAT v6.00) The first six functional images were discarded for calibration of the signal, leaving 594 functional images for preprocessing. Motion correction was done using linear registration to the middle volume using MCFLIRT (FMRIB's Linear Registration Tool) and a spatial smoothing Gaussian Kernel of 3mm at full-width at half-maximum (FWHM) was applied. Anatomical images were brain extracted using the Brain Extraction Tool (Smith, 2002). Matrices for later registration of functional images were then registered to the respective subject's anatomical images using 6 DOF affine registration, and T1 images were registered to the MNI152 standard image using 12 DOF affine registration and non-linear registration (FNIRT). Transformation matrices were not applied to the data until higher-level analysis. Next, ICA-AROMA (Pruim et al., 2015) was used to remove secondary head movement components using non-aggressive denoising. Components were checked manually and corrected if necessary. Last, nuisance regression was applied to remove signal from cerebrospinal fluid and white matter regions using linear regression. A second 5.2 mm FWHM Gaussian Kernel was applied to reduce noise, and an additional highpass filter of 0.01 Hz was applied to the data to remove slow drifts.

First-level Analysis

Two separate whole-brain regression analyses were completed with both the individual accelerometry regressors (amplitude and first derivative) on either day one or day two with $p < .001$ uncorrected. We thus obtained two zstat-images per day, which reflect the correlation of the regressors with each voxel's time course.

Second-level Analysis

Standard-space registration matrices were applied to the zstat-images of all subjects using FEAT. We performed two independent mixed-effects group-level analysis, one for day one and another for day two, to identify the group average tremor amplitude-related activation

in the brain. Second-level results were used to confirm earlier results regarding tremor-amplitude-related brain activity in the motor cortex, ventrolateral thalamus, and cerebellum, and to guide our choice for ROI's. An FWE cluster-correction of $p < .05$ was used. Activity related to the first derivative regressor is outside the scope of this thesis.

Replicability of Clinical Rating Scale and Accelerometry

In the main literature, the intra-class-correlation coefficient (ICC) is widely used as a measure of test-retest, intrarater, and interrater reliability (Koo & Li, 2016). Here, we used this method to assess replicability of clinical- and ACC measurements. One can choose between an 'absolute agreement' or 'consistency' model. Absolute agreement concerns if the same subjects keep the exact same score, while consistency concerns if scores on day one correlate with the scores on day two (e.g. if scores on day two were consistently 1 point higher than on day one). We used the absolute agreement model for random effects (where both measurements and subjects are random), which is indicated as model ICC (2,1) by Shrout & Fleis (1979). For the purpose of the present study, ICCs were interpreted following the main literature; 0.00–0.50 = poor; 0.51–0.75 = moderate; 0.76–0.90 = good; 0.91-1.00 = excellent reliability (Koo & Li, 2016).

Unfortunately, an ICC is not directly applicable to the outcome measures that we selected for our fMRI data. Therefore, we will use another method to assess replicability of fMRI-related parameters. However, to keep consistency and comparability within this paper, we also apply this technique to the clinical- and accelerometry-data, additionally to the ICC. The technique is based on comparing the between- and within-subject absolute differences between outcome measures from both sessions, to assess whether selected outcome measures are significantly more similar within subjects than between subjects (i.e. replicable). For all outcome measures of clinical- and accelerometry data, within-subject differences were calculated by subtracting the data of day one from day two and taking the absolute value. This

is done for every subject and for every outcome parameter. Between-subject absolute differences were calculated by comparing data of a ‘reference’ subject to the data of every other subject on the same day. This is done for both days, resulting in 12 values per day, and 24 in total. We consider the mean of all reference-to-database combinations of one particular subject as the between-subject difference for this specific subject. Each subject is used as the ‘reference’ once, thus we calculated one within- and one between-subject absolute difference per outcome measure for every subject. These values are used for statistical testing described below. Figure 1 provides a visual representation of this method.

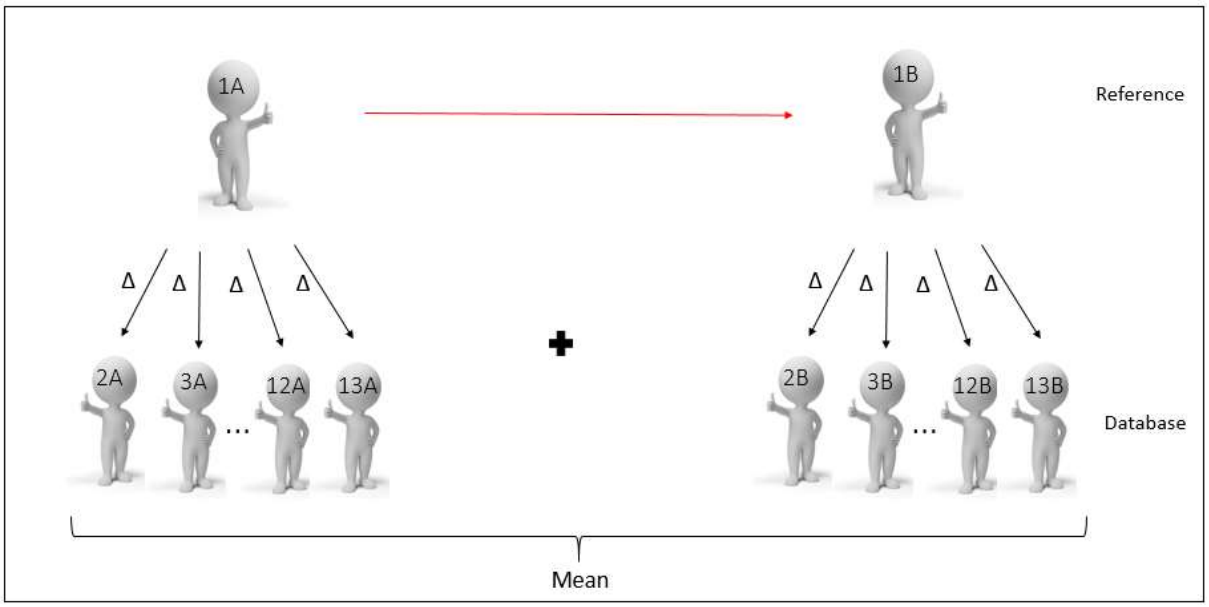

Figure 1. Visual representation of the method used for calculation of the within- and between-subject differences. Numbers represent subjects and letters represent days (A=day one, B=day two). The red line depicts within-subject comparisons. The black lines depict between-subject comparisons. Subject one on day one (1A) is compared to all other subjects on the same day, and subject one on day two (1B) is compared to all other subjects on the same day. We take the mean of all values and considered this as the between-subject difference. Every subject is the reference once.

Replicability of the CTC-network

Unthresholded amplitude-related zstat-images of every subject on both days were used to perform further analysis on using the fslcc tool (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012). We calculate the spatial correlation for each individual ROI (contralateral motor cortex, ventrolateral thalamus, ipsilateral cerebellum) and the primary visual cortex (V1) as a control region. The within-subject correlation is calculated by correlating the masked zstats of both sessions of one subject with each other. The between-subject correlation is calculated of each subject in a similar manner as described above for the between-subject differences for the tremor assessment (shown in Figure 1). We define a 'reference' image, which is a masked zstat of one subject on day one and correlate it with each zstat-image of other subjects on the same day. The same is done for day two and all these correlation values are averaged to obtain a single between-subject correlation value per subject. Thereby, we obtain two correlation values (one within- and one between-subject) per subject for each ROI separately. Statistical testing is performed on these values. Statistical tests, as described below, are applied to test if within-subject correlations are higher than between-subject correlations.

Replicability of the Peak Activation Location within each ROI of the CTC-network

The Euclidean distance between the peak activation locations of the amplitude-related zstats from both fMRI sessions of a subject is calculated for each ROI using the following formula:

$$d(\mathbf{p},\mathbf{q}) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2}.$$

Where p1 is the x-coordinate on day 1, and q1 is the x-coordinate on day 2, and so on for coordinates y and z. We use FSLstats (a command line function) to extract the coordinates (x-

y-z) of the maximum amplitude voxel per region of interest, because for clinical applications (such as deep brain stimulation) it might be relevant to known if the location of maximum tremor-related activation has a high within-subject stability. Thus, we extract the x-y-z coordinates of the contralateral motor cortex, the ventrolateral thalamus, ipsilateral cerebellum, and the primary visual cortex as a control region on day one and day two, resulting in eight locations per subject. We then calculate the within-subject distance between the peak activation locations of both sessions of a subject for all three ROI's separately. An average between-subject distance is calculated using the same method as shown in Figure 1. We use the peak activation location of one session of a subject as a reference and calculate the distance within the peak activation locations of every other subject on the same day. This is done for both days and all 24 values are averaged to obtain a single between-subject value per subject. We thus obtain a single within- and between-subject distance which we use for statistical testing described in the next paragraph.

Statistical Analysis

The Shapiro-Wilk test indicated that the between- and within-subject differences-, correlations- and distances were not normally distributed for any outcome parameter ($p < .05$ for all parameters). Therefore, we used the Mann-Whitney-U-Test, which is a non-parametric test for independent measures. We hypothesize that within-subject absolute differences are smaller than between-subject absolute differences for clinical-, and accelerometry data. We also hypothesize that within-subject distances are smaller than between-subject distances and that within-subject correlations are higher than between-subject correlations, since this would indicate replicability and uniqueness of outcome parameters. Results were considered statistically significant when one-tailed $p < .05$.

Results

Patients

We included 13 patients (10 men, three women) in the analysis. All 13 patients were measured on two days with on average 78 days in between (*SD*=51). Six patients had a left-hand tremor. Their MRI data was flipped using fslswapdim to make data comparable to patients with a right-hand tremor. Additional patient characteristics are presented in table 1.

Table 1.

Subject characteristics

Characteristic	**Mean (*SD*)**
Age in years	61 (5.0)
Disease duration in years	6.2 (2.4)
Disease score (UPDRS-III)	24 (7.8)
Male gender (%)	10 (77)
Days between measurements	78 (51)
FAB score	15 (1.8)
MOCA score	27 (2.6)

Note. Mean disease characteristics of all 13 patients are shown. Disease duration was defined as the time in years since official diagnosis. Days between measurements refers to the time delay between the first and second measurement.

Clinical group characteristics

Mean clinical rest tremor scores are presented in Table 2, for both day 1 and day 2 separately. Clinical rest tremor scores did not differ significantly between day 1 and day 2, as indicated by a paired-samples t-test (t_{12} =1.00, p = .337). All comparisons between days were two-tailed, as no assumptions of directionality were made. Significance was reached when $p < .05$.

Accelerometry group characteristics

Mean values for frequency and amplitude in both conditions are presented in Table 2, for both day 1 and day 2 separately. Neither rest frequency (t_{12} = 1.47, p = .165), coco frequency (t_{12} = -0.52, p = .610), rest amplitude (t_{12} = 1.46, p = .170) or sensitivity to cognitive stress (t_{12}= 1.42, p = .178) differed significantly between days as indicated by a paired-samples t-test. Significance was reached when $p < .05$.

Table 2.

Group results for clinical- and accelerometry data

Characteristic	Day 1	Day 2	Sig (two tailed, $p<0.05$)
TRS			
Rest tremor	2.61 (0.50)	2.54 (0.51)	$p = .337$
ACC			
Rest frequency	4.65 (0.70)	4.50 (0.78)	$p = .165$
Rest amplitude	22.27 (3.91)	21.70 (3.92)	$p = .170$
Coco frequency	4.90 (0.56)	4.96 (0.73)	$p = .610$
Coco amplitude	2.24 (2.34)	1.76 (2.06)	$p = .178$

Note. Mean values are shown with *SD* between parenthesis. Values for TRS and ACC frequencies are raw values, while power values are log transformed. The power in coco is relative to the rest condition, indicating sensitivity to cognitive stress.

Replicability of individual clinical- and accelerometry characteristics

The ICC for clinical resting tremor scores is .85 (CI=.70-.95, $p < .001$), which indicates *good* replicability between days.

Two ACC-related outcome variables show an ICC > .90, which indicates *excellent* replicability**.** For rest frequency, we find an ICC of .95 (CI = .85-.98, $p = .004$) and for rest power we find an ICC of .93 (CI = 0.79-0.97, $p = .003$). Frequency in the coco condition has an ICC of .88 (CI = .68-.96, $p = < .001$) and power in the coco condition shows an ICC of .84

(CI = .67-.94, $p < .001$), thus replicability is classified as *good*. Figure 2 is a visual representation of the findings.

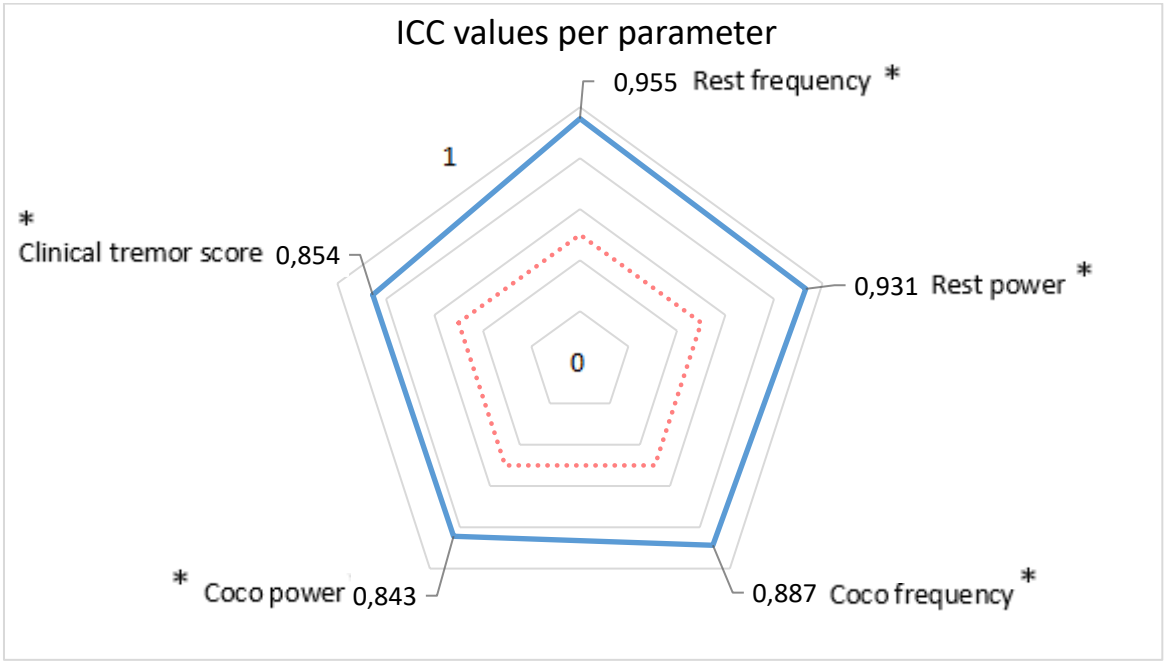

Figure 2. Intra-class correlation coefficients depicted for every parameter and connected through the blue line. All values are depicted on a scale of zero (inside) to one (outside), with steps of .2 in between. The red line indicates .5; ICC's below this value are considered not replicable. The asterisk (*) indicates significant effects when tested against $p < .05$.

For consistency purposes, we also computed within- and between-subject differences and performed statistical analysis on these variables. As expected, we find that the within-subject difference is significantly smaller than the between-subject difference for clinical rest tremor score ($U = 13$, $p = < .001$). We find the same effect for rest tremor frequency ($U = 27$, $p = < .001$), coco tremor frequency ($U = 9.0$, $p = 0.004$) rest power ($U = 0.0$, $p = < .001$), coco power ($U = 24$, $p < .001$), indicating individual replicability and uniqueness for both clinical- and behavioral accelerometry parameters. All comparisons are one-tailed, as we hypothesize smaller within-subject differences. The results for the accelerometry data are shown in figure 3.

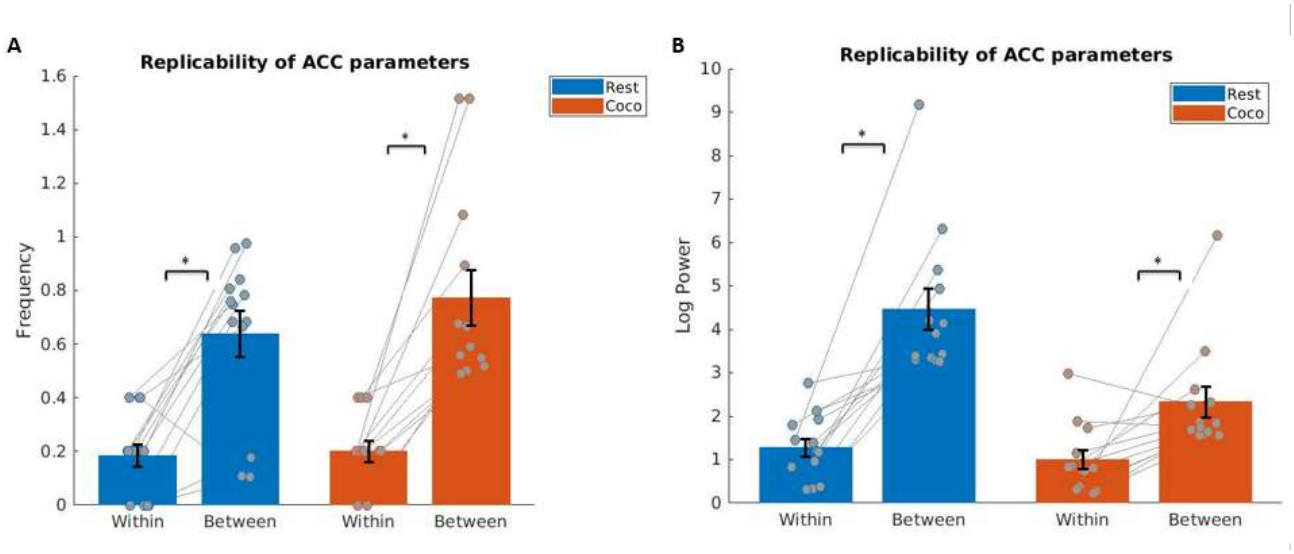

*Figure 3***.** The left panel (A) indicates frequency differences on the y-axis, where the blue bars indicate rest condition and the orange bars indicate coco condition. Within the conditions, we compared the within- to the between-subject differences. The right panel (B) shows the same but with differences in log transformed power values on the y-axis. As indicated by statistical testing, the within-subject differences are smaller than the between-subject differences for all comparisons.

Group level CTC-network results

We used concurrent accelerometry-fMRI to identify the tremor-amplitude related network. Masks for the contralateral motor cortex and the ipsilateral cerebellum were used based on the findings of Helmich *et al.,* (2011). The mask of the ventrolateral thalamus is based on the paper by Dirkx *et al.,* (2017). Group effects for both days separately are shown in figure 4.

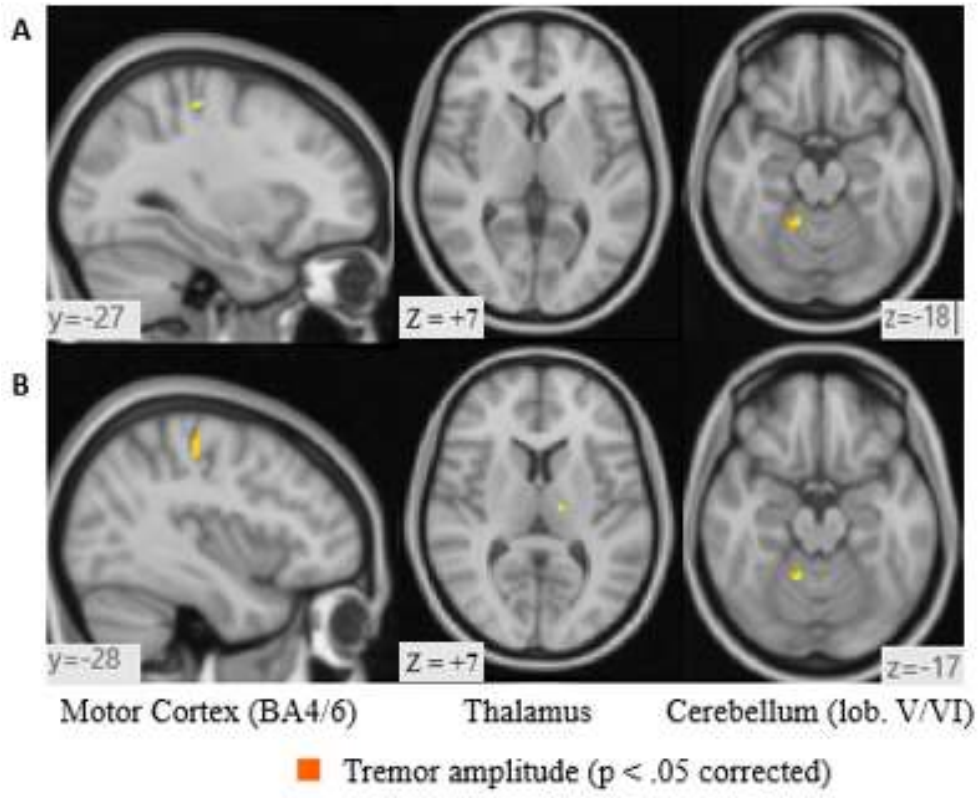

Figure 4. Group level results of two independent masked regression analyses with both the amplitude and first derivative regressor in the model. Figure shows areas that correlate with tremor amplitude-related activity (z-contrast on the amplitude regressor shown at a threshold of $p < .05$ using an intensity threshold of $Z > 3.1$). The right-side is the side contralateral to the tremor.

Panel A displays tremor-related activity on day one in the contralateral motor cortex (p=.025, 9 voxels) and the ipsilateral cerebellum ($p < .001$, 18 voxels). Panel B displays significant tremor-related activity on day two, localized to the motor cortex ($p < .001$, 56 voxels), and the cerebellum ($p < .001$, 38 voxels). The contralateral ventrolateral thalamus shows significant tremor-related activity on day two only ($p = .002$, 7 voxels).

Replicability of CTC-network

Using fslcc, we tested the stability of the spatial pattern of activation within the tremor-amplitude related network: the contralateral motor cortex, the ventrolateral thalamus, and the ipsilateral cerebellum. Besides our ROI's, we used the primary visual cortex as a control region to help interpretation. We expect the within-subject correlation to be higher than the between-subject correlation for the regions of the CTC-circuit.

We do not find significant differences in correlation values for the motor cortex (U = 71, p = .25) the cerebellum (U = 72, p = .27), or the thalamus (U = 57, p = .08). For the primary visual cortex, we also find no significant differences in spatial correlation values within- versus between-subjects (U = 66, p = .18). Figure 5 is a visual representation of the results.

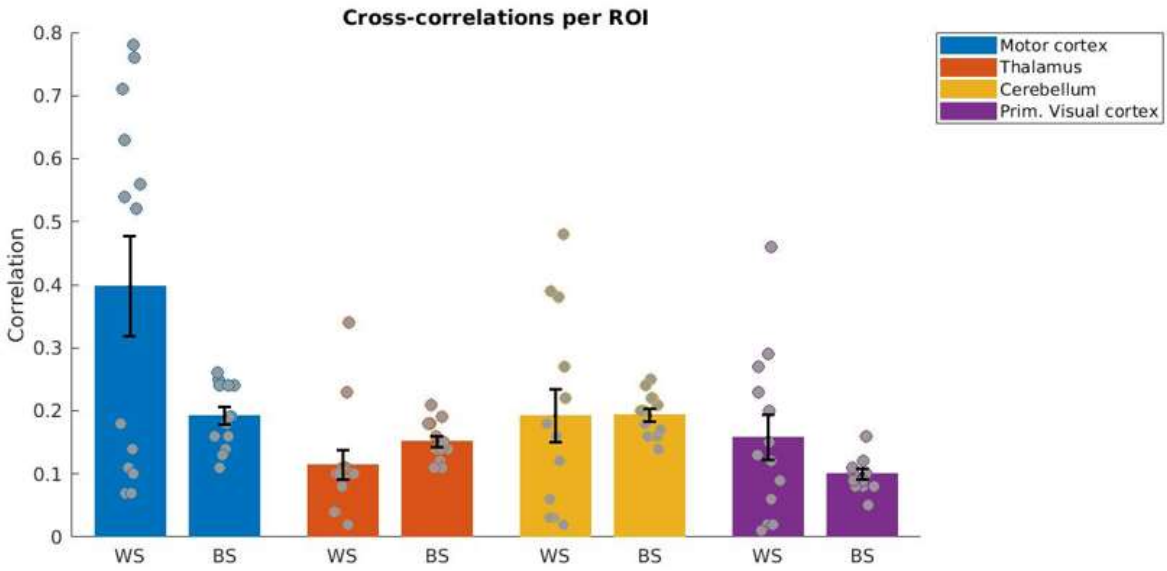

Figure 5. Within- and between-subject correlations in the motor cortex, the ventrolateral thalamus, the ipsilateral cerebellum, and the primary visual cortex. WS=within-subject; BS=between-subject.

Replicability of the Peak Activation Location

Between and within-subject Euclidean distances were calculated for each unilateral ROI in the tremor-amplitude related network and for the visual cortex as a control region. We expect the within-subject distance to be smaller than the between-subject distance, since this would indicate replicability and uniqueness. We find no significant differences when comparing within- and between-subject distances, nor for the motor cortex (U =74, p =.30), the ventrolateral thalamus (U =59, p =.19), and the cerebellum (U =61, p =.24). We find significantly smaller within- than between-subject distances for the primary visual cortex (U = 6, p < .001) Figure 6 is a visual representation of the results.

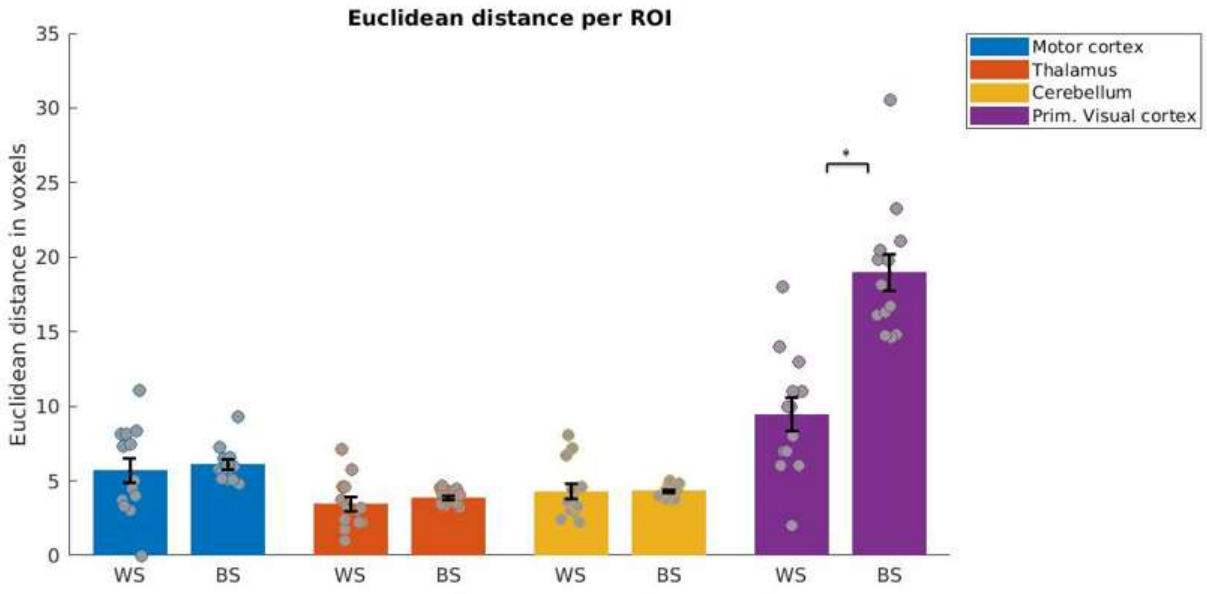

Figure 6. Within (WS)- and between-subject (BS) distances in the motor cortex, the ventrolateral thalamus, the ipsilateral cerebellum and the primary visual cortex. Distance is depicted in voxels.

Discussion

We investigated whether tremor-related characteristics are replicable (indicating stability over time) and unique to an individual (i.e. differ between subjects). Our analysis of clinical- and behavioral tremor characteristics revealed significantly smaller differences within- then between subjects for all parameters, which indicates that it is unique to an individual and stable over time.

We found ICC values ranging from .80-.97, with rest frequency as most stable, followed by rest amplitude. Previous studies support that the group-average tremor frequency is relatively stable over time, as indicated by a longitudinal study by Hellwig et al (2009), where they measured patients once per year for two- to four consecutive years. They find that tremor-frequency on average decreases with 0.09 Hz per year and thus remains relatively stable. Tremor amplitude is found to be more variable over longer time courses. Thielgen *et al.*, (2004) show data of a 10-hour EMG measurement of two patients clearly showing diurnal variations in tremor amplitude. Amplitude has further been assessed using a wearable accelerometry sensor over a 24-hour period where daily variation in amplitude was found over a group of 25 PD-patients (Smeja et al., 1999). They also find variance in the amplitude-timeseries between-subjects (i.e. one patient where tremor amplitude was highest during the night while for another patient it was highest between 2-4am) and within-subject (i.e. the overall amplitude of one patient was decreased by 15% three weeks later while the pattern of daily variation stayed relatively stable). Our study might not observe these diurnal variations because we only measure tremor for a small amount of time instead of over the day, thereby decreasing the effect of spontaneous diurnal variations. Moreover, during our study we measure tremor in a standardized environment, while the latter studies record tremor during daily activities. This might explain why we find more stable results, and it would indicate that variability of tremor amplitude is possibly context- and time-of-day-dependent. Furthermore,

it is striking that our results indicate that amplitude in rest is more replicable than frequency and amplitude during cognitive-coactivation. A previous study by Zach *et al* (2017) investigated the effect of cognitive-coactivation on the variability of tremor amplitude in the same standardized environment using the same trial time. They found that cognitive-coactivation significantly decreased tremor variability (indicated by a lower coefficient of variation) when compared to rest tremor, which would point towards higher replicability of coco-amplitude. We thus additionally examined the coefficient of variation (COV) of our rest and cognitive-coactivation data to test for the hypothesis of a ceiling-effect, i.e. the hypothesis that tremor during coco increases to roughly the same level for all participants, thus decreasing the uniqueness and decreasing the ICC. Our data indicates no significant difference in COV between rest and cognitive co-activation over days, thus this would not explain our results. Our finding of relatively lower replicability within the cognitive-coactivation condition also cannot be explained by subjects being more susceptible to cognitive stress on either day, since we find no differences between days regarding sensitivity to cognitive-coactivation. Additionally, one might argue that, besides the cognitive stress induced by our design, there are other session-specific stressors (e.g. private issues, reduced sleep) that could have influenced tremor amplitude**,** thereby inducing differences between days, thus increasing variance and decreasing the ICC. However, since we would expect these factors to influence all parameters within one day equally, it cannot explain why coco-related characteristics are relatively less replicable than rest-parameters. Thus, it is striking that both cognitive co-activation and rest condition show good-to-excellent replicability and uniqueness, but it should be taken into account that our study only incorporated 13 subjects, increasing the possibility of finding false-positive results. We conclude that it could be relevant to use these parameters as part of a *fingerprint* that might be used to differentiate patients into different groups (e.g. different speed of disease progression) and in the future to

predict possible treatments. However, first it is important that our findings are reproduced using a larger sample size. If our findings are reliably found in independent cohorts, future studies might focus on identifying possible underlying groups of patients.

Our group-level fMRI results indicate that the CTC-circuit is involved in tremor amplitude, as earlier established by Helmich *et al.* (2011). We find significant activation in the motor cortex and cerebellum on both days, and for the thalamus on day two only. Although relevance of the CTC-network in tremor-related activity is independently established in our cohort, it is striking that we find relatively better group-level results for day two (e.g. thalamic activation and larger clusters). We tested for a difference in movement, but the relative movement during the fMRI scan did not differ significantly between days, thus this cannot explain this finding.

We tested two fMRI parameters for uniqueness and stability. Individual tremor activation maps did not show significantly higher spatial correlation within-subjects than between-subjects for any of our ROI's. It is important to note that the variance in the within-subject correlations is relatively high for the motor cortex, indicating a bimodal pattern in the distribution of the correlations. About half of the subjects show high within-subject correlational values (> .6), while the other half shows extremely low correlational values (< .2), while no subject shows a correlation between 0.2 and 0.6. This might indicate that there are underlying factors that influence our findings by producing a bimodal pattern, specifically in the motor cortex.

Furthermore, we do not observe significant difference in within- versus between-subject distances between peak activations in either of the ROI's, but we do find significantly lower within-subject distances in the visual cortex when compared to between-subject distances. This implies that the location of the peak activation in V1 is replicable and unique, while the amplitude within each ROI of the CTC-network is not. We tested this parameter

because of its possible informative value for clinical practice, such as finding the optimal location for placement of deep-brain stimulation electrodes or focused ultrasound. A similar bimodal pattern as in spatial correlations is observed in the results for Euclidean distances of peak activation from the two sessions. We find either relatively high within-subject distances in the motor cortex (> 7.5 voxels/15 mm), or relatively low (< 5 voxels/10 mm) within-subject distances. Thus, this again indicates the probability of different underlying groups of patients. Moreover, our results, especially the low replicability of peak activation location, indicate that this parameter will most likely not be informative for physicians in making decisions about DBS and focused ultrasound targeting locations. For example, there is an average variance in DBS location sites between subjects of two millimeters in the subthalamic nucleus (Klein *et al*, 2012). Since the average variance in location is much smaller than the distances in the location of the maximum amplitude, our parameter will probably not be informative. Nevertheless, methodological confounds could have played a role. Again it is possible that inaccuracies of registration cause deviations in the location of the peak activation. Moreover, subjects could have had a different positioning of their arms during the two fMRI sessions. For example, minor changes in the support of the lower arms (e.g. by the pillows) cause have altered the tremor-movement, which would likely influence the location of the activation. Additionally, tremor severity might be different between sessions by means of daily variation, which could also possibly influence the nature of the tremor-movement. Amplitude-related parameters as used in our study might not be the optimal parameters for fingerprinting. Future studies might focus on identifying possible fingerprint characteristics of non-amplitude related fMRI parameters, such as effective connectivity between different nodes of the CTC-network.

Interpretational Issues

It should be mentioned that, on day two, almost all measurements took place at a later time of day (about 1.5 h), so diurnal variations (i.e. variations due to the time of the day) and medication effects (i.e. passed time since the last medication intake) cannot be excluded. However, there was no increase in tremor amplitude between days, which indicates that there was no major effect of time or medication intake. Also, all patients received 16-minutes of transcranial alternating current stimulation (outside the scanner) at individual tremor-frequency over the motor cortex on day two. This was always at least 30 minutes before both the accelerometry- and accelerometry-fMRI measurement. Research on the after-effects of tACS indicate that after 15-minute stimulation between 5-20 Hz, the duration of the after-effect (e.g. excitability of the motor cortex) is around 10 minutes (Schutter and Hortensius, 2011). Therefore, we do not expect the stimulation to have an effect on our findings.

Conclusion

In this study we identified clinical- and accelerometry parameters, such as frequency and amplitude, as highly replicable between sessions. We additionally find that these parameters are unique to an individual. Therefore, these parameters might be suitable for *fingerprinting,* i.e. identifying different groups of patients based on these characteristics. A larger follow-up study is required to assess the reliability of our findings. Moreover, we find no significant differences for spatial correlation and Euclidean distances in our ROI's. Although results for the Euclidean distances and spatial correlations in the motor cortex show a trend towards significance, their informative value for clinical practice seems limited. Moreover, since our study is based on 13 patients, we have to acknowledge that our small sample might reduce the reliability of our results.

References

Bezard, E., Dovero, S., & Prunier, C. (2001). Relationship between the appearance of symptoms and the level of nigrostriatal degeneration in a progressive 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-lesioned macaque model of Parkinson's disease. *Journal of Neuroscience, 21,* 6853-6861.

Budzianowska, A., & Honczarenko, K. (2008). Assessment of rest tremor in Parkinson's disease. *Journal of Neurological Surgery, 42,* 12–21.

Deuschl, G., Bain, P., & Brin, M. (1989). Consensus statement of the Movement Disorder Society on Tremor. Ad Hoc Scientific Committee. *Movement Disorders, 13,* 2-23.

Dirkx, M.F., den Ouden, H.E., Aarts, E., Timmer, M.H., Bloem, B.R., Toni, I, & Helmich, R.C. (2017). Dopamine controls Parkinson's tremor by inhibiting the cerebellar thalamus. *Brain, 140*(3), 721-734.

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X.., & Constable, R. T. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature neuroscience, 18*(11), 1664-71.

Hellwig, B., Mund, P., Schelter, B., Guschlbauer, B., Timmer, J., & Lücking, C.H. (2009). A longitudinal study of tremor frequencies in Parkinson's disease and essential tremor. *Clinical Neurophysiology, 120*(2), 431-435.

Fahn S, Tolosa E, & Marín C. (1993) Clinical rating scale for tremor. In: Jankovic J, Tolosa E, editors. Parkinson's disease and movement disorders. 2nd ed. Baltimore: Williams & Wilkins; pp. 225–234.

Fearnley, J.M., Lees, A.J. (1991). Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain, 114*(5), 2283-2301.

Helmich, R.C., Janssen, M.J., Oyen, W.J., Bloem, B.R., & Toni, I. (2011). Pallidal dysfunction drives a cerebellothalamic circuit into Parkinson tremor. *Annals of Neurology, 69*(2), 269-281.

Helmich, R.C., Hallet, M., Deuschl, G., Toni, I., & Bloem, B.R. (2012). Cerebral causes and consequences of parkinsonian resting tremor: a tale of two circuits? *Brain, 135*, 3206-3226.

Kish, S.J., Shannak, K., & Hornykiewicz, O. (1988). Uneven Pattern of Dopamine Loss in the Striatum of Patients with Idiopathic Parkinson's Disease. *New England Journal of Medicine, 318*, 876-880.

Klein, J.C, Barbe, M.T., Seifried, S., Baudrexel, M. Runge., M., Maarouf, M., Gasser, E., Hattingen, T., Liebig, R., Deichmann, L., Timmermann, L., Weise, R. (2012). The tremor network targeted by successful VIM deep brain stimulation in humans. *Neurology*, *78*(11), 787-795.

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, *15*(2), 155-63.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., & Smith, S.M. (2012). FSL. *Neuroimage, 62*, 782:790.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2010). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *201*, 156-869.

Pruim, R.H.R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., & Beckmann, C.F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artefacts from fMRI Data. *Neuroimage, 112*, 267-277.

Schutter D. J. L. G., Hortensius R. (2011). Brain oscillations and frequency-dependent modulation of cortical excitability, *Brain Stimulation*, *4.* 97–103.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bullletin, *78*, 420–428.

Smeja, M., Foerster, F., Fuchs, G., Emmans, D., Hornig, A., & Fahrenberg, J. (1999). 24-h assessment of tremor activity and posture in Parkinson's disease by multi-channel accelerometry. *Journal of Psychophysiology*, *13*(4), 245-256.

Smith, S.M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*. *17*(3), 143-155

Thielgen, T., Foerster, F., Fuchs, G., Hornig, A., & Fahrenberg, J. 2004. Tremor in Parkinson's disease: 24-hr monitoring with calibrated accelerometry. *Electromyography in Clinical Neurophysiology*, *44*(3).

Waller, L., Walter, H., Kruschwitz, J.D., Reuter, L., Muller, S., & Erk, S. (2017). Evaluating the replicability, specificity, and generalizability of connectome fingerprints. *Neuroimage, 58*, 371-377.

Zach, H., Dirkx, M., Bloem, B. R., & Helmich, R. C. (2015). The Clinical Evaluation of Parkinson's Tremor. *Journal of Parkinson's disease*, *5*(3), 471-474.

Zach, H., Dirkx, M. F., Pasman, J. W., Bloem, B. R., & Helmich, R. C. (2017). Cognitive Stress Reduces the Effect of Levodopa on Parkinson's Resting Tremor. *CNS Neuroscience & Therapeutics*, *23*(3), 209-215.