Journal of the Neurological Sciences 397 (2019) 16–21

Contents lists available at ScienceDirect

Journal of the Neurological Sciences

journal homepage: www.elsevier.com/locate/jns

A baseline study for detection of Parkinson's disease with 3D-transcranial sonography and uni-lateral reconstruction

Annika Plate[a,*], Juliana Maiostre[a], Johannes Levin[a], Kai Bötzel[a], Seyed-Ahmad Ahmadi[a,b]

[a] Department of Neurology, Ludwig-Maximilians-University, Marchioninistr. 15, 81667 Munich, Germany
[b] German Center for Vertigo and Balance Disorders, Ludwig-Maximilians-University, Munich, Germany

ARTICLE INFO

Keywords:
Parkinson's disease
Transcranial sonography
3D ultrasound

ABSTRACT

Introduction: TCS is a well-established technique for diagnosis of Parkinson's disease (PD). Volumetric 3D-TCS is a promising complementary approach for objective acquisition and analysis, in particular for less experienced sonographers. This study provides baselines for Parkinson detection (sensitivity and specificity), cutoff values and inter-rater agreement in 3D-TCS.

Methods: We performed 3D-TCS in 52 subjects (healthy controls and PD) bilaterally, and reconstructed in 3D space uni-laterally. Ipsi-lateral hyperechogenicities in the substantia nigra are manually segmented slice-by-slice in the 3D volume by two raters at different experience levels. ROC threshold analysis is performed and compared on features representing 3D volume and axial cross-sections (2.5D) of hyperechogenicities. Pearson correlation and intra-class correlation coefficients were evaluated for assessment of inter-rater agreement.

Results: 50 subjects were included. Both raters achieved high classification accuracy with 2.5D/3D features extracted from 3D-TCS volumes (best results sensitivity/specificity/cut-off per rater: 84.6%/88.9%/25.0mm^2; 77.8%/88.9%/95.9mm^3). The inter-rater agreement in 3D was high (ICC(A,1) = 0.777, $p < 10^{-3}$), the classification performance of both sonographers was statistically not significantly different.

Conclusion: The study presents first baseline values for uni-lateral 3D-TCS examination, and finds no disadvantage of uni-lateral reconstructions compared to previous bi-lateral fusion. Volumetric 3D-TCS has potential for a high inter-rater agreement and accuracy in detection of PD, in particular for sonographers with less experience.

1. Introduction

Two-dimensional (2D) transcranial B-mode sonography (TCS) has become a supporting tool for the diagnosis of Parkinson's disease (PD) and other diseases ever since the first studies [1]. It has been shown in several studies that an enlarged hyperechogenicity of the substantia nigra area (SN) is a relevant finding in PD patients [2–4]. By now it has been established as a level A recommendation in the new European guidelines for the diagnosis of PD [5]. This is also due to this method being easily accessible, non-invasive and providing a cost-efficient approach.

In a meta-analysis from 2016, the TCS technique itself presented a high diagnostic accuracy with a pooled sensitivity of 83% and specificity of 87% in detecting PD [6]. Another recent meta-analysis of 2017 detected a prevalence rate of hyperechogenicity in the SN of 84% in idiopathic Parkinson's disease (IPD), with a 75% sensitivity and 70% specificity to differentiate IPD from atypical parkinsonism [7]. Therefore, the TCS technique has proven reliable in providing useful and supporting information for the detection of IPD.

Importantly, a longitudinal multi-center study on 1847 older subjects without PD at baseline has shown that SN hyperechogenicities indicate a 17.4 times higher risk of developing PD within a 37 months follow-up [4] and a 20.6 times higher risk after a five year follow-up [8]. Early diagnosis and intervention in PD may help preserve neuron function and consequently improve quality-of-life by reducing later symptoms, while reducing overall treatment costs for patients [9].

As the EFNS-MDS guidelines state [5], analysis of midbrain and SN with two-dimensional TCS is feasible and reliable, if sonographers have enough experience and if adjustment to consensus guidelines are guaranteed. Experience is necessary for choosing the optimal axial level and image plane tilt during the examination, as well as for the correct outlining of hyperechogenicities in the SN, in particular in more

* Corresponding author at: Ludwig-Maximillians-University, Department of Neurology, Marchioninistr. 15, 81667 Munich, Germany.
E-mail addresses: Annika.Plate@med.uni-muenchen.de (A. Plate), Johannes.Levin@med.uni-muenchen.de (J. Levin), Boetzel@med.uni-muenchen.de (K. Bötzel), Ahmad.Ahmadi@med.uni-muenchen.de (S.-A. Ahmadi).

https://doi.org/10.1016/j.jns.2018.12.001
Received 16 September 2018; Received in revised form 20 November 2018; Accepted 1 December 2018
Available online 03 December 2018
0022-510X/

difficult cases of bone windows with increased acoustic impedance and beam scattering. Consequently, the 2D method is recommended to be performed by trained readers [10]. There have been a few attempts at facilitating 2D-TCS, mostly through (semi-)automatic computerized analysis of 2D ultrasound images. For example, concerning the measurement of the SN, automatized methods with texture analysis [11], active contour algorithms [12] or user-interactive morphometric evaluation of SN features [13] have been proposed.

While these methods help with analysis of the acquired 2D TCS images, the challenge and subjectivity of optimal TCS image plane alignment through the SN remain. To this end, a series of works from our group have explored three-dimensional (3D) TCS as a novel imaging paradigm, since it maps the whole mesencephalic region into one volume, thus removing the need for optimal cut-plane selection and leading to higher repeatability and objectivity of the image acquisition. As shown in earlier pilot studies of our group [14–16], first attempts of differentiating PD from healthy controls with a bi-lateral 3D-TCS reconstruction technique revealed promising diagnostic results with a sensitivity and specificity of up to 90.9 and 72.7%.

Our pilot approach [16] required bi-lateral reconstruction from both bone windows into one fused volume, which is technically difficult and error-prone. Here, we set the first expert-derived baseline for differentiation of PD patients from healthy controls (HC) in a uni-lateral 3D reconstruction scheme. Classification accuracy and inter-rater variability from two raters with different levels of experience under this paradigm were evaluated. The cohort size in this study provides sufficient statistical power, allowing us to substantiate and expand upon previously made claims regarding 3D-TCS towards diagnosis of IPD.

2. Material and methods

2.1. Clinical data

52 subjects (healthy controls, $N = 25$, and PD patients, $N = 27$) were recruited from the outpatient clinic of the department of neurology. General inclusion criteria were an age of 18 or above and no history of brain surgery. PD subjects were diagnosed with PD according to the UK Parkinson's Disease Society Brain Bank [17]. Age-matched healthy controls (HC) were partners or friends of PD patients. Inclusion criteria for the HC subjects were a negative medical history for neurological diseases and the lack of abnormalities in the neurological examination. Additional information was obtained (age, sex, Unified Parkinson's Disease Rating Scale (UPDRS) part III, Hoehn&Yahr (H&Y) and disease duration). The study was approved by the local ethics committee. All subjects gave their written informed consent. Further details on the cohorts can be found in Table 1, respectively.

2.2. Acquisition and reconstruction of 3D TCS

Volumetric acquisition was performed bi-laterally with a 3D Freehand ultrasound setup, similar to our earlier study [16]. A clinical ultrasound machine (Ultrasonix MDP, BK Ultrasound, Peabody, MA, USA) was used for sonographic imaging, with a transducer suitable for transcranial ultrasound (center frequency range 2.5–5.0 MHz). A custom software was developed for real-time acquisition and reconstruction of B-mode ultrasound data.

Next to the ultrasound machine, an optical tracking system (NDI Polaris Spectra; Northern Digital Inc., Waterloo, ON, Canada) was used for tracking of 3D poses (six degrees of freedom) of two tracking targets. The first target was attached to the transcranial transducer, for real-time acquisition of a 3D + time tracking stream. Combining 3D transducer poses with the 2D + time image stream allowed for reconstruction of 3D image volumes. Reconstruction was performed uni-laterally, creating a volume for each sweep at 0.5 mm isotropic voxel resolution with a backward compounding algorithm (Gaussian distance kernel, voxel range 2.0) [18]. The second target was attached to the forehead of patients to avoid reconstruction artifacts due to unconstrained head motions. Reconstruction artifacts were further contained through accurate calibration of the 3D-TCS setup, a slow manual transducer motion at high acquisition rates (25 Hz for TCS images and 60 Hz for tracking data) and the backward compounding algorithm, which together helped in reducing e.g. tear and gap artifacts.

For each subject, two to five 3D-TCS sweeps were recorded through the left and right bone windows for subsequent uni-lateral 3D reconstruction. These sweeps were performed by positioning the transducer on the pre-auricular bone window and then, from a midbrain cross-section position, the image plane was tilted downwards in caudal direction until the brainstem, where 3D acquisition was started. Acquisition of a single 3D sweep took approximately 30 s.

2.3. Segmentation of ultrasound data

Two investigators, blinded to patient diagnosis throughout the study and with different experience levels in TCS sonography (A: medium, P: high), agreed on one volume for the left and right bone windows respectively, to provide common grounds for the segmentation and data analysis. Next, each rater individually outlined hyperechogenicities of the SN in multiple slices manually. The first slice underneath the third ventricle was chosen as a cranial start-point, and segmentation sustained 5.5 mm further caudal, resulting in 11 slices (0.5 mm per slice).

Table 1
Details for subjects from Healthy Controls (HC) and Parkinson's Diseases (PD) groups (IQR: inter-quartile range). Significance p-values are either not computed (n.c.) based on statistical tests: Fisher's exact test for binary variables (FET) and Wilcoxon rank-sum test (WRS) for non-normal distributed variables (normality rejected by Shapiro-Wilk, significance level $p < .05$; p-values in the last six rows reported after Bonferroni-Holm multiple testing correction). PD types are akinetic-rigid (AR), tremor-dominant (TD) and mixed-type (M).

	HC		PD		Level of significance
	Median / IQR	Range	Median / IQR	Range	
Sex	13 female / 10 male	–	11 female / 16 male	–	p: 0.395 (FET)
Age	66.0 / 4.5	53–78	71.0 / 7.8	41–82	p: 0.065 (WRS)
PD Dominant Side	–	–	12 right / 11 left / 4 none	–	n.c.
PD Type	–	–	12 AR / 11 TD / 4 M	–	n.c.
PD Hoehn&Yahr	–	–	2.0 / 0.9	1–4	n.c.
PD Duration [years]	–	–	9.0 / 7.1	2–20	n.c.
UPDRS	–	–	29.0 / 19.2	10–66	n.c.
Vol. Midbrain [mm3]	2504.7 / 519.3	346.0–3327.8	2473.5 / 563.5	1809.1–3658.2	p: 0.59611 (WRS)
Vol. SN Left [mm3]	52.0 / 46.7	7.5–142.1	120.6 / 70.1	7.0–319.3	p: 0.00041 (WRS)
Vol. SN Right [mm3]	47.1 / 34.9	4.1–153.1	99.8 / 108.2	13.6–285.2	p: 0.00015 (WRS)
Feature: 2.5D max [mm2]	20.3 / 11.3	0.0–39.8	45.3 / 21.0	6.4–75.9	p < .00001 (WRS)
Feature: 3D max [mm3]	55.9 / 48.5	0.0–153.1	150.9 / 85.3	13.6–319.3	p < .00001 (WRS)
Feature: 3D sum [mm3]	84.3 / 78.4	0.0–275.3	230.4 / 182.1	18.9–502.8	p: 0.00012 (WRS)

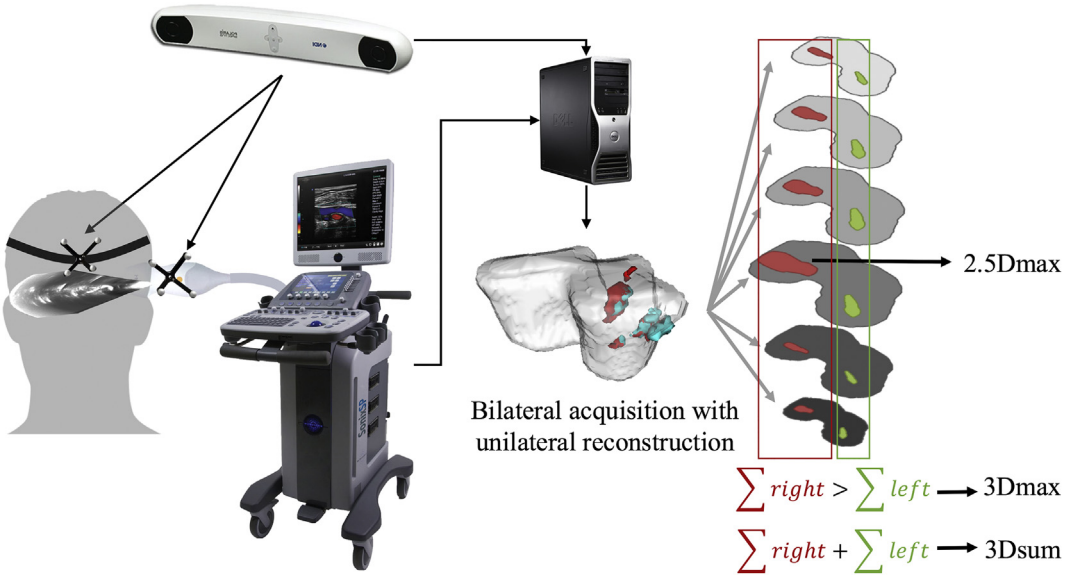

Fig. 1. Graphical abstract of 3D TCS setup and methodology: 2D images from an ultrasound machine are simultaneously acquired for each bone window, with calibrated 3D tracking information of the transducer pose. Both data streams are combined in a workstation for 3D uni-lateral reconstruction of volumes. Analysis is performed bi-laterally (volumes from left and right bone window): ipsi-lateral SN hyperechogenicities are manually segmented, and features 2.5Dmax, 3Dmax and 3Dsum are extracted for group comparison.

Selection of the first slice within the volume and optimal contrast settings for segmentation were decided upon independently by each rater. For each volume, manual segmentation took approximately 15–25 min using the software ITK Snap [19]. Three different regions of interests were considered: ipsilateral hyperechogenicities in the area of the SN on both sides and the outline of the midbrain. For segmentation, the auto-contrast setting of the imaging software was manually adjusted for each sweep, due to large variability of image appearance across subjects and bone windows (see Fig. 1).

2.4. ROC classification and parameters

For characterization of PD patients and healthy controls (HC), we performed binary classification with a receiver-operator-characteristic (ROC) approach [3], maximizing the cost-effective cutoff of a single continuous parameter that represents the amount of SN hyperechogenicity in each subject. We performed ROC classification of single-dimensional features (area or volume) extracted from left and right SN hyperechogenicities, to determine the optimal cutoff for prospective separation of HC vs. PD. The 3D-TCS derived features, illustrated in Fig. 1, are defined as:

"2.5D max": largest area (in mm^2) of SN hyperechogenicity in a single slice of the 3D SN volume of either side in one subject. The axial location of the maximum does not necessarily have to be localized on the same level for each sweep or subject.

"3D max": largest 3D SN volume (in mm^3), i.e. the larger unilateral volume of hyperechogenic SN 3D-regions from both hemispheres.

"3D sum": sum of 3D SN volumes (in mm^3), i.e. the sum of 3D-segmented SN echogenicity volumes in the left and right hemisphere

2.5. Statistics

We determined the necessary cohort size a priori using a power analysis [20] based on "2.5D max" feature values of a single expert rater obtained in our previous pilot study [14]. For a statistical power of 0.90, we calculated the required cohort size for this study to be at least $N = 46$ (one-tailed t-test, allocation ratio N1/N2 = 1.0, group HC $25.6 +/- 12.6 mm^3$, group HC $35.2 +/- 9.0 mm^3$, effect size d = 0.877, significance level $\alpha = 0.05$, power $1-\beta = 0.90$).

Differences of median values in features "2.5D max", "3D max" and "3D sum" between the groups HC and PD were tested using a non-parametric Wilcoxon rank-sum test, after rejecting a normal distribution of the data (Shapiro-Wilk test, significance level $p < .05$). Multiple testing outcomes were compensated using a stringent Bonferroni-Holm correction for familywise error rate. We also calculated the linear correlation of the three 3D-TCS derived measures "2.5D max", "3D max" and "3D sum" with the UPDRS scores (Pearson's rho) and the Hoehn& Yahr (Spearman's rho) outcomes in the PD group.

ROC analyses and comparisons were performed using Matlab (The Mathworks, Inc.) and Statistical Analysis of ROC curves (StAR) [21]. In addition to ROC classification accuracy for two raters separately, we also analyzed the inter-rater agreement. For better assessment of inter-rater variation, we report results from both raters separately, in order of increasing examiner experience. We calculated inter-rater agreement in terms of Pearson's correlation coefficient, as well as intraclass correlation coefficient (ICC) according to [22], which accounts for random-chance agreement between raters. We calculate ICC(A,1), i.e. the absolute agreement between both raters, modeled as random effects, with a single measure per rater [23].

3. Results

3.1. Qualitative 3D TCS results

Before quantitative analysis, we show axial cross-sections of reconstructed TCS volumes for three different patients with different bone window quality in Fig. 2. Notable features are the midbrain region, the opposite cranial wall and the difference in tissue insonification of the ipsi- and contra-lateral brain hemispheres. In particular, the visibility of contra-lateral tissue structures such as lateral ventricle boundaries strongly depends on the bone window properties. The lower degree of contra-lateral structural visibility is also visible inside the midbrain, justifying the approach of solely segmenting the ipsi-lateral midbrain ROI, similar to the 2D approach [2,24]. In Fig. 2, it is also noteworthy that the initial auto-contrast appearance of image volumes set by the imaging software (left two columns) is different than after a manual contrast adjustment (column three), where images appear more

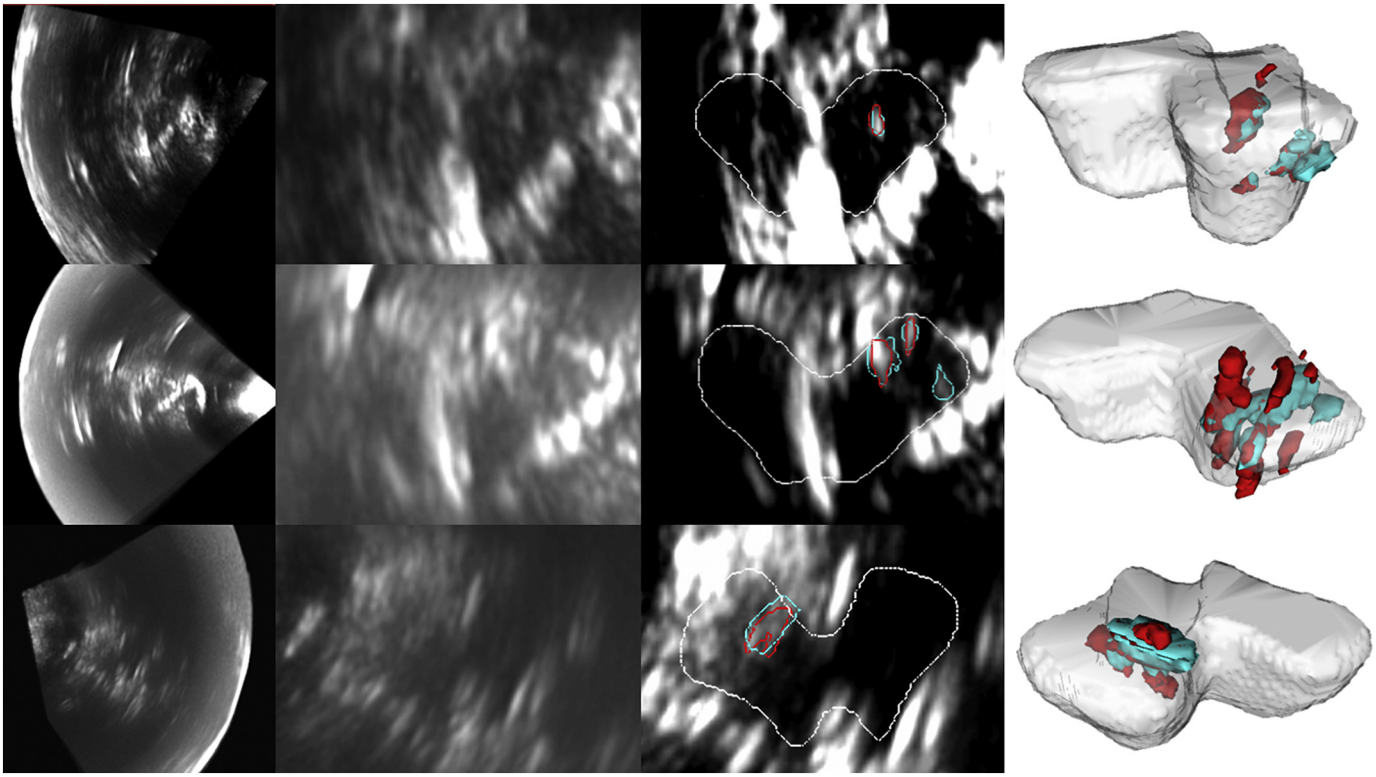

Fig. 2. Three uni-lateral example 3D TCS scans from different subjects (top two rows scanned from left bone window, bottom row scanned from right), with decreasing bone window quality from top to bottom. The first column shows an axial cross-section through the whole-brain, the second column an enlargement of the midbrain section (auto-contrast set by software). Column three shows manually adjusted image contrast for manual segmentation. White outline shows the midbrain shape, and colored outlines the segmentations of ipsi-lateral SN Hyperechogenicities by both raters (cyan: AA, red: AP). The fourth column shows a 3D visualization of midbrain surface and 3D SN hyperechogenicity surfaces by both raters (cyan: AA, red: AP). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

homogeneously. Outlines of SN hyperechogenicity segmentations for both raters (color-coded, cyan A, red P) exhibit overlap, but not necessarily for all SN speckle patches, and not necessarily in all slices (cf. 3D view, column four).

3.2. Quantitative sonographic data

Data acquisition was performed on 52 subjects. Insufficient bone windows for TCS examination were observed in seven subjects (13.5%). Of those, two subjects (3.9%) showed bi-laterally insufficient bone windows and had to be excluded from the study. Bone windows of the five other subjects (9.6%) allowed for at least uni-lateral sufficient TCS examination of the SN, enabling inclusion of those subjects in the study (overall inclusion rate: 96.2%). The final cohort contained 23 healthy control subjects and 27 IPD subjects. Details on the cohort regarding demographics and IPD scores can be found in Table 1, along with statistical significance of group differences for each variable.

The difference between the HC cohort and the PD group was most obvious in the 2.5D (HC: 20.3mm^2, IQR: 11.3 versus PD: 45.3mm^2, IQR: 21.0) and the 3D max. measurement method (HC: 55.9mm^3, IQR: 48.5 versus PD: 150.9mm^3, IQR: 85.3) which showed a high significance level ($p < .00001$). Median volumetric SN hyperechogenicities of the HC in the left and right hemisphere (52.0 mm^3 and 47.1 mm^3, respectively) was less than half the size compared to the PD cohort (120.6 mm^3 and 99.8mm^3).

There is no linear relationship between clinical outcome measures and overall amount of observed hyperechogenicity derived from 3D-TCS, as measured by “2.5D max”, “3D max” and “3D sum” (for UPDRS: all rho < 0.04, $p > .85$; for Hoehn&Yahr: all rho < 0.13, $p > .52$).

3.3. ROC classification results

In one rater (AA), best classification results according to AUC value (with 95% confidence intervals) were achieved with the “3D max” feature, with a sensitivity of 77.8% and a specificity of 88.9%, at a cut-of value of 95.9mm^3.

For the other rater (AP), best results were seen with the “2.5D max” method reaching 84.6% sensitivity and a specificity of 88.9%, at a cut-off value of 25.0 mm^2. A post-hoc power analysis of all volumetric measurement techniques (2.5D-max, 3D max. and 3D sum) yielded a high statistical power ($\beta < 0.0001$) with 2.5 max. and 3D max. yielding highest statistical power (see Table 1).

Detailed results of sensitivity, specificity, cut-off values as well as AUC values with 95% confidence intervals for both raters and the three features are given in Table 2. The AUC values of all 3D-approaches ranged between 0.82 and 0.87 in both raters, with the feature “2.5D max” performing best in terms of AUC.

The difference of ROC classification outcomes between both raters and the three features was further tested towards statistical significance using Statistical Analysis of ROC curves (StAR) [21]. Regarding difference of both raters, we found no statistically significant difference in ROC outcomes, neither for features “2.5D max” (p: 0.531), nor for feature “3D max” (p: 0.226), or “3D sum” (p: 0.515). Regarding difference of features, the ROC comparison yielded no statistically significant result in any rater, given our cohort (rater P: “2.5D max” vs. “3D max” with p: 0.297, “2.5D max” vs. “3D sum” with p: 0.788, “3D max” vs. “3D sum” with p: 0.756; rater A: “2.5D max” vs. “3D max” with p: 0.479, “2.5D max” vs. “3D sum” with p: 0.273, “3D max” vs. “3D sum” with p: 0.158).

Table 2
Classification results with a single-dimensional feature and a single threshold chosen from a ROC analysis (AUC: area under the curve, CI: 95% confidence interval).

Single-dimensional Feature (AP)	Single Threshold with ROC			
	Sensitivity (CI) (%)	Specificity (CI) (%)	AUC (CI)	Threshold
2.5D max: max. of both largest areas	84.6 (80.0–88.7)	88.9 (84.6–92.5)	0.87 (0.73–0.95)	25.0mm^2
3D max: larger volume side (11 slices)	84.6 (80.0–88.7)	85.2 (80.6–89.2)	0.82 (0.71–0.94)	76.9 mm^3
3D sum: volume of both sides (11slices)	80.8 (75.9–85.2)	85.2 (80.6–89.2)	0.83 (0.71–0.94)	116.9 mm^3
2.5D max: max. of both largest areas	77.8 (72.8–82.3)	85.2 (80.6–89.2)	0.85 (0.74–0.95)	27.3 mm^2
3D max: larger volume side (11 slices)	77.8 (72.8–82.3)	88.9 (84.7–92.5)	0.85 (0.74–0.96)	95.9 mm^3
3D sum: volume of both sides (11slices)	77.8 (72.8–82.3)	92.6 (88.7–95.7)	0.83 (0.71–0.94)	154.5mm^3

3.4. Inter-rater agreement results

The median difference between both raters in selecting the most cranial slice for manual segmentation was 0.5 mm (min: 0.0 mm; first quartile: 0.0 mm; third quartile: 1.5 mm; max: 4.5 mm). Volumetric 3D measurement of ipsi-lateral SN yielded a high Pearson correlation between both raters of 0.798 ($p < 10^{-3}$), with high intra-class correlation ICC(A,1) of 0.777 ($p < 10^{-3}$).

4. Discussion

In this study, we analyzed 3D TCS data of 50 subjects including HC and PD subjects. Towards a quantitative and reliable way of classifying between those two groups, we present a 3D-TCS acquisition method of analyzing volumetric segmentation of the midbrain and hyperechogenicities in the area of the SN by analyzing data uni-laterally.

In general, classification accuracy and inter-rater agreement on our data were was high when using volumetric calculations, either as 2.5D or 3D. ROC classification was accurate yielding AUC values consistently above 0.82, and did not show statistically significant differences across features or raters. A recent meta-analysis on the 2D method [6] yielded a pooled analysis with 83% sensitivity and specificity 87% to distinguish PD from HC with TCS. Sensitivity in the analyzed studies ranges from 63%–100%, specificity from 31%–97%. Our 3D results are in general comparable to these pooled results.

Differences in ROC cutoff values between both raters appear widely varied in absolute values, but considering the total ranges of each feature across both raters, the relative differences are very comparable ("2.5D max": 0.67%, "3D max": 0.65%, "3D sum": 0.87%). Together with the non-significant differences in ROC results, we conclude that the three investigated features have comparable performance and outcomes. Overall, the 3D method yielded consistent results across both raters, and across three different volumetric features, starting with a consistent choice of the most-cranial slice during the manual segmentation process (only 0.5 mm or one slice median difference between raters). Together, this is supportive towards our hypothesis that volumetric analysis is robust and promotes objectivity. The fact that the sonographer with less experience approached the classification outcome of the expert, who served as the gold standard, is noteworthy. Contributing factors might be volumetric analysis of the whole SN, and averaging of several hundred images into a volume, which leads to an improved signal-to-noise ratio and increased robustness of segmentation in 3D.

In terms of 3D imaging technology, we relied on a unilateral reconstruction, which is novel and different from our first 3D pilot study [14], where we used a bilateral reconstruction approach. In this study, we show that this does not lead to disadvantages, but may actually be beneficial. Exclusion rates due to insufficient bone windows in this study (3.9%) did not increase compared to bi-lateral reconstruction in [14] (4.3%), and 2D TCS (10–20%) [25,26]. At the same time, unilateral reconstruction is technically easier to achieve, since the bi-lateral reconstruction used in the former study is dependent on a highly accurate tracking technology and calibration, both of which are challenging to achieve. Uni-lateral reconstruction allows for using entirely trackerless setups, including 2D matrix-array ultrasound machines and transducers [27], handheld ultrasound imaging devices with integrated gyroscope [28] or inference of 3D transducer pose from imaging data directly using artificial intelligence [29]. Together with open-source software frameworks for 3D freehand ultrasound [30], immediate exploration of the 3D method by other research groups and clinics is now much more feasible.

There are some study limitations. In this study, we intentionally left out an in-depth analysis of intra-rater agreement and decided to focus instead on an analysis of inter-rater agreement, which is typically larger in magnitude. Still, an analysis of intra-rater variability would be interesting to investigate, but due to a much higher effort for manual segmentation in 3D-TCS compared to 2D slices, this needs to be pursued in future work. A second limitation stems from freely adaptable brightness and contrast settings during manual segmentation by each rater. Different contrast settings can have an effect on segmentation outcome, and may vary across raters. For future studies, we recommend setting a fixed contrast, e.g. by winsorization of image intensities within an appropriate image ROI (e.g. within the midbrain). A further limitation of the 3D technique so far is the need for manual segmentation. While we needed to perform manual segmentation to provide gold standard volumetric measures in this study, a manual analysis of 3D-TCS is still too time-consuming for clinical routine. In future work, automatic analysis techniques are required to cut down processing time and introduce further objectivity. The values for classification outcome and inter-rater variability in this study can serve as baselines for evaluation of automated methods in future work.

5. Conclusion

In this study, we present the first quantitative analysis of bi-lateral SN hyperechogenicities in 3D-TCS given uni-lateral reconstructions. For this setup, we set first baselines for volumetric measurements of midbrain and SN segmentation by two raters. Our cohort comprised healthy controls and PD patients and had sufficient statistical power to re-new and substantiate claims made in our pilot study on 3D-TCS [14] regarding classification accuracy and cutoff values of 2.5D and 3D features. Compared to bi-lateral reconstructions, a uni-lateral 3D-TCS reconstruction approach is technically easier to achieve, and provides comparable classification accuracy, inter-rater objectivity and ease of use, especially for more inexperienced sonographers.

Full Financial Disclosure/Conflicts of interests:

Dr. Annika Plate: reports no disclosures.

Juliana Maiostre: reports no disclosures.

Dr. Johannes Levin: Consultancies: Axon Neuroscience, Ionis Pharmaceuticals; Honoraria: Bayer, Vital; other: Abbvie.

Prof. Dr. K. Boetzel: Honoraria: Medtronic.

Dr. Seyed-Ahmad Ahmadi: reports no disclosures

Funding

This work was funded by the Lüneburg Heritage, by Deutsche Forschungsgesellschaft (DFG) grant BO 1895/4–1, as well as the German Federal Ministry of Education and Research (BMBF) in connection with the foundation of the German Center for Vertigo and Balance Disorders (DSGZ) (grant number 01 EO 0901).

The funding sources had no involvement in the study design, collection, analysis, interpretation of the data or writing or publication of the article.

References

[1] G. Becker, J. Seufert, U. Bogdahn, H. Reichmann, K. Reiners, Degeneration of substantia nigra in chronic Parkinson's disease visualized by transcranial color-coded real-time sonography, Neurology 45 (1995) 182–184, https://doi.org/10.1212/WNL.45.1.182.

[2] U. Walter, D. Dressler, T. Probst, A. Wolters, M. Abu-Mugheisib, M. Wittstock, R. Benecke, Transcranial brain sonography findings in discriminating between parkinsonism and idiopathic Parkinson disease, Arch. Neurol. 64 (2007) 1635.

[3] S. Van De Loo, U. Walter, S. Behnke, J. Hagenah, M. Lorenz, M. Sitzer, R. Hilker, D. Berg, Reproducibility and diagnostic accuracy of substantia nigra sonography for the diagnosis of Parkinson' s disease, (2010), pp. 1087–1093, https://doi.org/10.1136/jnnp.2009.196352.

[4] D. Berg, K. Seppi, S. Behnke, I. Liepelt, K. Schweitzer, H. Stockner, F. Wollenweber, A. Gaenslen, P. Mahlknecht, J. Spiegel, J. Godau, H. Huber, K. Srulijes, S. Kiechl, M. Bentele, A. Gasperi, T. Schubert, T. Hiry, M. Probst, V. Schneider, J. Klenk, M. Sawires, J. Willeit, W. Maetzler, K. Fassbender, T. Gasser, W. Poewe, Enlarged substantia nigra hyperechogenicity and risk for parkinson disease: a 37-month 3-center study of 1847 older persons. Arch. Neurol. 68 (7) (2011) 932–937, https://doi.org/10.1001/archneurol.2011.141.

[5] A. Berardelli, G.K. Wenning, A. Antonini, D. Berg, B.R. Bloem, V. Bonifati, D. Brooks, D.J. Burn, C. Colosimo, A. Fanciulli, J. Ferreira, T. Gasser, F. Grandas, P. Kanovsky, V. Kostic, J. Kulisevsky, W. Oertel, W. Poewe, J.-P. Reese, M. Relja, E. Ruzicka, A. Schrag, K. Seppi, P. Taba, M. Vidailhet, EFNS/MDS-ES/ENS [corrected] recommendations for the diagnosis of Parkinson's disease., Eur, J. Neurol. 20 (2013) 16–34, https://doi.org/10.1111/ene.12022.

[6] D.-H. Li, Y.-C. He, J. Liu, S.-D. Chen, Diagnostic Accuracy of Transcranial Sonography of the Substantia Nigra in Parkinson's disease: A Systematic Review and Meta-analysis, Sci. Rep. 6 (2016) 20863, , https://doi.org/10.1038/srep20863.

[7] A. Shafieesabet, S.M. Fereshtehnejad, A. Shafieesabet, A. Delbari, H.R. Baradaran, R.B. Postuma, J. Lokk, Hyperechogenicity of substantia nigra for differential diagnosis of Parkinson's disease: A meta-analysis, Park. Relat. Disord. 42 (2017) 1–11.

[8] D. Berg, S. Behnke, K. Seppi, J. Godau, S. Lerche, P. Mahlknecht, I. Liepelt-Scarfone, C. Pausch, N. Schneider, A. Gaenslen, K. Brockmann, K. Srulijes, H. Huber, I. Wurster, H. Stockner, S. Kiechl, J. Willeit, A. Gasperi, K. Fassbender, T. Gasser, W. Poewe, Enlarged hyperechogenic substantia nigra as a risk marker for Parkinson's disease, Mov. Disord. 28 (2013) 216–219, https://doi.org/10.1002/mds.25192.

[9] D.L. Murman, Early treatment of Parkinson's disease: opportunities for managed care, Am. J. Manag. Care 18 (2012) S183–S188.

[10] P. Prati, A. Bignamini, L. Coppo, A. Naldi, C. Comi, R. Cantello, G. Gusmaroli, U. Walter, The measuring of substantia nigra hyperechogenicity in an Italian cohort of Parkinson disease patients: a case/control study (NOBIS Study), J. Neural Transm. 124 (2017) 869–879, https://doi.org/10.1007/s00702-017-1724-9.

[11] L. Chen, J. Hagenah, A. Mertins, Feature analysis for Parkinson's disease detection based on transcranial sonography image, Med. Image Comput. Comput. Interv. – MICCAI 2012, 2012, pp. 272–279.

[12] A. Sakalauskas, A. Lukoševičius, K. Laučkaite, D. Jegelevičius, S. Rutkauskas, Automated segmentation of transcranial sonographic images in the diagnostics of Parkinson's disease, Ultrasonics 53 (2013) 111–121, https://doi.org/10.1016/j.ultras.2012.04.005.

[13] J. Blahuta, T. Soukup, M. Jelinkova, P. Bartova, P. Cermak, R. Herzig, D. Skoloudik, A new program for highly reproducible automatic evaluation of the substantia nigra from transcranial sonographic images, Biomed. Pap. 158 (2014) 621–627, https://doi.org/10.5507/bp.2013.029.

[14] A. Plate, S.A. Ahmadi, O. Pauly, T. Klein, N. Navab, K. Bötzel, Three-dimensional sonographic examination of the midbrain for computer-aided diagnosis of movement disorders, Ultrasound Med. Biol. 38 (2012) 2041–2050, https://doi.org/10.1016/j.ultrasmedbio.2012.07.017.

[15] S.A. Ahmadi, M. Baust, A. Karamalis, A. Plate, K. Boetzel, T. Klein, N. Navab, Midbrain segmentation in transcranial 3D ultrasound for Parkinson diagnosis, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) (2011) 362–369.

[16] O. Pauly, S.-A. Ahmadi, A. Plate, K. Boetzel, N. Navab, Detection of substantia nigra echogenicities in 3D transcranial ultrasound for early diagnosis of Parkinson disease, Med. Image Comput. Comput. Assist. Interv. 15 (2012) 443–450.

[17] A.J. Hughes, S.E. Daniel, L. Kilford, A.J. Lees, Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases, J. Neurol. Neurosurg. Psychiatry 55 (1992) 181–184.

[18] W. Wein, F. Pache, B. Röper, N. Navab, Backward-warping ultrasound reconstruction for improving diagnostic value and registration, Med. Image Comput. Comput. Assist. Interv. 9 (2006) 750–757.

[19] P.A. Yushkevich, J. Piven, H.C. Hazlett, R.G. Smith, S. Ho, J.C. Gee, G. Gerig, User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability, NeuroImage 31 (2006) 1116–1128, https://doi.org/10.1016/j.neuroimage.2006.01.015.

[20] F. Faul, E. Erdfelder, A. Buchner, A.G. Lang, Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses, Behav. Res. Methods (2009), https://doi.org/10.3758/BRM.41.4.1149.

[21] I.A. Vergara, T. Norambuena, E. Ferrada, A.W. Slater, F. Melo, StAR: a simple tool for the statistical comparison of ROC curves, BMC Bioinformatics. (2008), https://doi.org/10.1186/1471-2105-9-265.

[22] P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, Psychol. Bull. 86 (1979) 420–428.

[23] K.O. McGraw, S.P. Wong, Forming inferences about some intraclass correlation coefficients, Psychol. Methods 1 (1996) 30–46.

[24] D. Berg, Marker for a preclinical diagnosis of Parkinson'S disease as a basis for neuroprotection, (2006), pp. 123–132.

[25] A. Pilotto, R. Yilmaz, D. Berg, Developments in the role of transcranial sonography for the differential diagnosis of parkinsonism, Curr. Neurol. Neurosci. Rep. 15 (2015), https://doi.org/10.1007/s11910-015-0566-9.

[26] R. Yilmaz, D. Berg, Substantia Nigra Hyperechogenicity for the Early Detection of Parkinson's Disease, J. Neurol. Neurophysiol. 08 (2017).

[27] M. Ivancevich, G. Pinton, S.W. Smith, H.A. Nicoletto, M. Scism, E. Bennett, D.T. Laskowitz, 7A-5 Real-Time 3D Contrast-Enhanced Transcranial Ultrasound, Ultrason. Symp. (2007), https://doi.org/10.1109/ULTSYM.2007.143.

[28] F.J.N. Mancuso, V.N. Siqueira, V.A. Moisés, A.F.T. Gois, A.A.V. de Paola, A.C.C. Carvalho, O. Campos, Focused cardiac ultrasound using a pocket-size device in the emergency room, Arq. Bras. Cardiol. (2014), https://doi.org/10.5935/abc.20140158.

[29] Raphael Prevost, Mehrdad Salehi, Julian Sprung, Alexander Ladikos, Robert Bauer, Wolfgang Wein, Deep learning for sensorless 3D freehand ultrasound imaging, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2017, pp. 628–636, , https://doi.org/10.1007/978-3-319-66185-8_71.

[30] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, G. Fichtinger, PLUS: Open-source toolkit for ultrasound-guided intervention systems, IEEE Trans. Biomed. Eng. 61 (2014) 2527–2537, https://doi.org/10.1109/TBME.2014.2322864.