

# Association of obstructive sleep apnea with brain volumetry and cognition in de novo Parkinson's disease

Josef Mana

2024-08-12

## Methods

### Neuropsychological assessment

All patients (PD) and healthy controls (HC) were administered a battery of neuropsychological tests including screening via Montreal cognitive assessment (MoCA) (Kopecek et al. 2017; Nasreddine et al. 2005), and assessment of (i) declarative memory via Rey Auditory Verbal Learing Test (RAVLT) (Bezdicek et al. 2014; Frydrychová et al. 2018); (ii) attention via Trail Making Test, part A (TMT-A) (Bezdicek et al. 2012; Bezdicek, Stepankova, et al. 2017), and dot colour naming (PST-D) as well as naming colour of neutral words (PST-W) conditions from Prague Stroop Test (Bezdicek et al. 2015); (iii) executive function via Trail Making Test, part B (Bezdicek et al. 2012; Bezdicek, Stepankova, et al. 2017), and Prague Stroop Test, interference condition (i.e., naming colour of contrasting colour words, PST-C) (Bezdicek et al. 2015); and (iv) processing speed via Grooved Pegboard Test (GPT) (Kløve 1963). The patients were further examined using tests from the standard International Parkinson and Movement Disorder Society (MDS) neuropsychological battery at Level II for mild cognitive impairment in Parkinson's disease (PD-MCI) (Litvan et al. 2012; Bezdicek, Sulc, et al. 2017). The Czech normative calculator established by Bezdicek, Sulc, et al. (2017) was used to assign PD-MCI diagnosis to each PD patient.

### Statistical analyses

Demographic (age, education, gender) and descriptive clinical (RBD, MoCA, BMI, BDI-II, age at first motor symptom, disease duration, MDS-UPDRS I, II and III) variables were described by their mean and standard deviation if continuous and frequency if nominal separately for HC OSA-, HC OSA+, PD OSA-, and PD OSA+ groups of participants, and compared by Gaussian (continuous variables) or logistic (binary variables) regression with group (PD vs. HC), OSA (OSA+ vs. OSA-) and their interaction as predictors if variables were measured in both PD and HC groups, and with OSA only as predictor if variables were measured in PD group only (i.e., disease-specific variables). As we did not aim to control type I error rate in these analyses, decision threshold for claiming statistically significant difference in

demographic and descriptive clinical variables was set at  $p < .05$  without adjustment for multiple comparisons.

### Cortical thickness

### Subcortical volumetry

Distribution of subcortical structures' volume conditional on group (PD vs HC) and OSA (OSA+ vs OSA-) was evaluated by a set of univariate linear regressions fitted to data via QR decomposition. Each subcortical structure's volume was regressed on group, OSA and their interaction as exposures of interest, and years of age, gender and TIV as additive linear covariates (i.e., no covariate/exposure interactions were allowed). All continuous variables were standardised (i.e., mean-centred and scaled by their in-sample standard deviation) before entering the analysis, and nominal variables were entered via sum coding. Each regression was tested for deviation from normality of residuals via the Shapiro-Wilk test and deviation from homoscedasticity via the Breusch-Pagan test. Based on the results of the primary analysis outlined above, we further explored association between hippocampal fields, and group and OSA in a series of post-hoc univariate regressions with identical specification as described above. The only difference between the primary analysis of the gross subcortical structures described above, and the analysis of hippocampal fields was that in the latter case, we focused on the group/OSA interaction only. The decision threshold for rejecting the hypothesis of zero group/OSA interaction was set according to adjustment for 5% FDR within this set of analyses.

### Cognitive variables

Distribution of cognitive performance indexes conditional on group and OSA was evaluated by a set of univariate linear regressions with identical specification as described in the Subcortical volumetry section with the only difference being that TIV was not included in the covariate set but education was. Moreover, response time variables were log-transformed and then reversed (i.e., multiplied by negative one) before standardising. Finally, since it is generally accepted that the variance increases with the mean in response time variables (Wagenmakers and Brown 2007), we re-fitted relevant cognitive models (i.e., models with statistically significant difference in mean outcome conditional on any of the exposures) with statistically significant deviation from homoscedasticity via a Bayesian distributional linear regression that allows for addressing heteroscedasticity by modelling it directly (Bürkner 2017). The re-fitted models included all the same terms as the original ones with added computation of residual standard deviation conditional on group, OSA, group/OSA interaction, age, sex, and education. The models were fitted using Stan's (version 2.21.0) build-in Hamiltonian Monte Carlo sampler accessed via R version 4.3.3 using package "brms" (Bürkner 2017; R Core Team 2024; Stan Development Team 2020)

with manually specified weakly informative priors and default “brms” settings otherwise (see the R code at [https://github.com/josefmana/pd\\_ippOSA.git](https://github.com/josefmana/pd_ippOSA.git)). As a check of results robustness to heteroscedasticity, parameter estimates of the Bayesian distributional models were then compared to the coefficients of the classical linear regressions.

### Statistical testing

Within the classical Neyman-Pearson hypothesis testing framework, we selected a set of decision thresholds for rejecting the hypothesis of zero difference between means of subcortical structure’s volume and cognitive indexes conditional on group/OSA by adjusting for 5% false discovery rate (FDR) (Benjamini and Hochberg 1995) in tests of all main effects and interactions of interest (i.e., coefficients related to the group and OSA variables) within each analysis. The test statistics consisted of  $t$ -values for each relevant regression coefficient calculated by `lm()` function in R software for statistical computing (R Core Team 2024). The primary estimand of interest, i.e., the interaction between group and OSA, was further characterised by calculating simple main effects of OSA within each group and the difference between these simple main effects via the `avg_comparisons()` function in the “marginaleffects” R package (Arel-Bundock, Greifer, and Heiss Forthcoming). In the exploratory analysis of hippocampal fields’ distributions, only the group/OSA interaction was considered and the decision threshold of this analysis was thus adjusted for 5% FDR for the number of hippocampal fields only.

Finally, to reflect current recommendations of the American Statistical Association which advised against basing scientific conclusions on whether a  $p$ -value passes a specific threshold (Wasserstein and Lazar 2016), we also calculated Shannon information (i.e.,  $s$ -values)  $-\log_2(p)$ .  $S$ -value is cognitive tool to help researchers intuitively evaluate strength of evidence against a null hypothesis contained in the results as equivalent to the number consecutive “heads” tosses that would provide the same amount of evidence against the null hypothesis that the coin is fair (Greenland 2019; Cole, Edwards, and Greenland 2021).

## Results

### Sample description

In total 45 PD patients (24 OSA- and 21 OSA+) and 41 HC subjects (21 OSA- and 20 OSA+) were included in the study. Descriptive statistics of the sample are summarised in Table 1. At the uncorrected 5%  $\alpha$  level, we reject hypotheses of zero difference between OSA+ and OSA- groups in mean BMI and percentage of men. More precisely, the OSA+ group in our sample was characterised by higher percentage of men and higher BMI compared to the OSA- group. There were neither any statistically significant differences in descriptive

**Table 1.** Demographic data of patients and healthy controls

	Descriptive statistics <sup>1</sup>				Inference statistics <sup>2</sup>		
	HC		PD		Group	OSA	Group * OSA
	OSA-	OSA+	OSA-	OSA+			
Sex (% male)	11 (52%)	18 (90%)	13 (54%)	14 (67%)	$z = -1.35$ , $p = .178$	$z = -2.47$ , $p = .013$	$z = 1.48$ , $p = .138$
RBD (% present)	-	-	9 (38%)	5 (24%)	-	$z = 0.98$ , $p = .326$	-
Age (years)	$60.58 \pm 8.69$	$61.12 \pm 8.31$	$60.25 \pm 8.60$	$62.94 \pm 8.50$	$t = 0.40$ , $p = .688$	$t = -0.87$ , $p = .384$	$t = -0.58$ , $p = .562$
Education (years)	$15.33 \pm 3.32$	$14.35 \pm 3.33$	$15.12 \pm 2.80$	$14.95 \pm 2.38$	$t = 0.31$ , $p = .760$	$t = 0.90$ , $p = .371$	$t = -0.63$ , $p = .530$
BMI	$25.55 \pm 3.44$	$28.63 \pm 3.62$	$26.91 \pm 3.03$	$29.72 \pm 3.87$	$t = 1.62$ , $p = .108$	$t = -3.91$ , $p < .001$	$t = 0.19$ , $p = .854$
Age at first symptom (years)	-	-	$59.07 \pm 9.13$	$61.63 \pm 8.54$	-	$t = -0.97$ , $p = .339$	-
Disease duration (years)	-	-	$1.38 \pm 0.66$	$1.29 \pm 0.49$	-	$t = 0.52$ , $p = .603$	-
MoCA (range 0-30)	$26.29 \pm 1.87$	$24.90 \pm 2.75$	$24.96 \pm 3.08$	$25.14 \pm 3.24$	$t = -0.89$ , $p = .374$	$t = 0.99$ , $p = .327$	$t = -1.29$ , $p = .199$
MDS-UPDRS I	-	-	$4.58 \pm 4.11$	$6.57 \pm 4.71$	-	$t = -1.51$ , $p = .137$	-
MDS-UPDRS II	-	-	$6.75 \pm 4.08$	$7.29 \pm 5.73$	-	$t = -0.36$ , $p = .717$	-
MDS-UPDRS III (total score)	-	-	$26.58 \pm 14.00$	$30.86 \pm 14.80$	-	$t = -0.99$ , $p = .325$	-
MDS-UPDRS III (axial subscore)	-	-	$6.25 \pm 3.98$	$6.14 \pm 3.45$	-	$t = 0.10$ , $p = .924$	-
MDS-UPDRS III (rigidity/akinesia subscore)	-	-	$15.21 \pm 8.74$	$18.33 \pm 10.40$	-	$t = -1.10$ , $p = .280$	-
MDS-UPDRS III (tremor subscore)	-	-	$5.12 \pm 3.96$	$6.38 \pm 4.86$	-	$t = -0.95$ , $p = .345$	-

<sup>1</sup>Presented as mean  $\pm$  standard deviation for continuous and count (percentage) for binary variables.<sup>2</sup>Based on Gaussian (continuous variables) or logistic (binary variables) regressions with OSA (disease-specific variables) or Group, OSA and Group \* OSA interaction predictor terms with coefficients computed via QR decomposition solution to the least squares problem as implemented in R `glm()` function.

HC, healthy controls; PD, patients with Parkinson's disease; OSA, obstructive sleep apnea; OSA+, group with moderate to severe OSA; OSA- group without moderate to severe OSA; Group, diagnosis group (PD vs HC); RBD, Rapid eye movement sleep Behavior Disorder; BMI, Body Mass Index; MoCA, Montreal Cognitive Assessment; MDS-UPDRS, Movement Disorders Society-Unified Parkinson's Disease Rating Scale.

variables between HC and PD groups nor any statistically significant group/OSA interactions.

## Cortical thickness

### Subcortical volumetry

Across linear regressions, models of left and right Accumbens showed deviation from homoscedasticity, and model of left Thalamus showed deviation from normality of residuals. None of the remaining models showed deviations from any assumption of linear regression (see Supplementary Table *subcortical\_regression\_coefficients.csv*). Figure 1 presents single data points organised by hemisphere, subcortical structure, diagnosis group and OSA. After adjusting decision threshold for 5% FDR, we can reject the null hypothesis of zero group/OSA interaction for age, sex and TIV adjusted hippocampal volume bilaterally (Table 2). This result stems from the observation that whereas in PD group, OSA+ subjects had reliably smaller hippocampal adjusted volumes than OSA- subjects, in HC group, the difference between OSA+ and OSA- subjects' hippocampal adjusted volumes did not reach statistical significance (Figure 2). Similar pattern of results was observed in Pallidum, Amygdala, and Caudate without reaching significance level at the adjusted 5% FDR.

### Hippocampal fields

Across linear regressions, only the model of left Presubiculum body showed deviation from normality of residuals, none of the remaining models showed any deviations from any assumption of linear regression (see Supplementary Table *hippocampi\_regression\_coefficients.csv*). The group/OSA interaction was detected using the 5% FDR adjusted threshold in all but right hippocampal tail adjusted volume in the longitudinal axis, in bilateral Presubiculum (both head and body), Subiculum body, and CA1 head as well as in right Parasubiculum, left Subiculum head, right CA4DG head and right HATA (see Table 3 and Figure 3).

**Table 2.** Results of regression analyses estimating means of subcortical structures' volume conditional on diagnosis group and obstructive sleep apnea status

	Left hemisphere							Right hemisphere						
	$\beta^I$	SE	95% CI	t value	p-value	s-value	sig.	$\beta^I$	SE	95% CI	t value	p-value	s-value	sig.
<b>Hippocampus</b>														
Group	-0.17	0.17	[-0.51, 0.17]	-0.990	.325	1.62		-0.20	0.18	[-0.55, 0.16]	-1.112	.270	1.89	
OSA	0.16	0.17	[-0.18, 0.49]	0.938	.351	1.51		0.18	0.17	[-0.16, 0.52]	1.057	.294	1.77	

**Table 2.** Results of regression analyses estimating means of subcortical structures' volume conditional on diagnosis group and obstructive sleep apnea status

	Left hemisphere							Right hemisphere						
	$\beta^I$	SE	95% CI	t value	p-value	s-value	sig.	$\beta^I$	SE	95% CI	t value	p-value	s-value	sig.
Group * OSA	1.16	0.32	[0.52, 1.79]	3.608	< .001	10.86	*	1.09	0.33	[0.44, 1.74]	3.319	.001	9.51	*
Amygdala														
Group	-0.07	0.17	[-0.42, 0.27]	-0.430	.668	0.58		-0.24	0.18	[-0.59, 0.11]	-1.363	.177	2.50	
OSA	-0.01	0.17	[-0.34, 0.32]	-0.049	.961	0.06		0.13	0.17	[-0.21, 0.47]	0.758	.451	1.15	
Group * OSA	0.51	0.32	[-0.12, 1.15]	1.606	.112	3.16		0.55	0.33	[-0.10, 1.21]	1.679	.097	3.37	
Thalamus														
Group	-0.32	0.15	[-0.62, -0.01]	-2.061	.043	4.55		-0.13	0.16	[-0.44, 0.18]	-0.855	.395	1.34	
OSA	0.08	0.15	[-0.21, 0.38]	0.564	.574	0.80		0.08	0.15	[-0.22, 0.38]	0.546	.587	0.77	
Group * OSA	-0.24	0.28	[-0.80, 0.33]	-0.836	.406	1.30		0.00	0.29	[-0.58, 0.58]	0.002	.998	0.00	
Caudate														
Group	0.06	0.21	[-0.36, 0.48]	0.281	.780	0.36		0.01	0.21	[-0.40, 0.42]	0.041	.967	0.05	
OSA	0.05	0.20	[-0.35, 0.46]	0.265	.792	0.34		-0.02	0.20	[-0.42, 0.38]	-0.098	.922	0.12	
Group * OSA	0.35	0.39	[-0.43, 1.13]	0.899	.371	1.43		0.60	0.38	[-0.16, 1.36]	1.573	.120	3.06	
Putamen														
Group	-0.19	0.17	[-0.53, 0.15]	-1.126	.263	1.92		-0.22	0.16	[-0.53, 0.10]	-1.374	.173	2.53	
OSA	0.01	0.17	[-0.32, 0.34]	0.060	.953	0.07		-0.14	0.15	[-0.44, 0.16]	-0.905	.368	1.44	
Group * OSA	-0.11	0.32	[-0.75, 0.52]	-0.355	.723	0.47		-0.18	0.29	[-0.76, 0.40]	-0.608	.545	0.88	
Pallidum														
Group	0.18	0.20	[-0.22, 0.58]	0.904	.369	1.44		0.00	0.20	[-0.40, 0.40]	0.010	.992	0.01	
OSA	0.40	0.19	[0.01, 0.78]	2.055	.043	4.53		0.28	0.19	[-0.10, 0.67]	1.452	.150	2.73	
Group * OSA	0.61	0.37	[-0.13, 1.34]	1.634	.106	3.23		0.77	0.37	[0.03, 1.51]	2.060	.043	4.55	

**Table 2.** Results of regression analyses estimating means of subcortical structures' volume conditional on diagnosis group and obstructive sleep apnea status

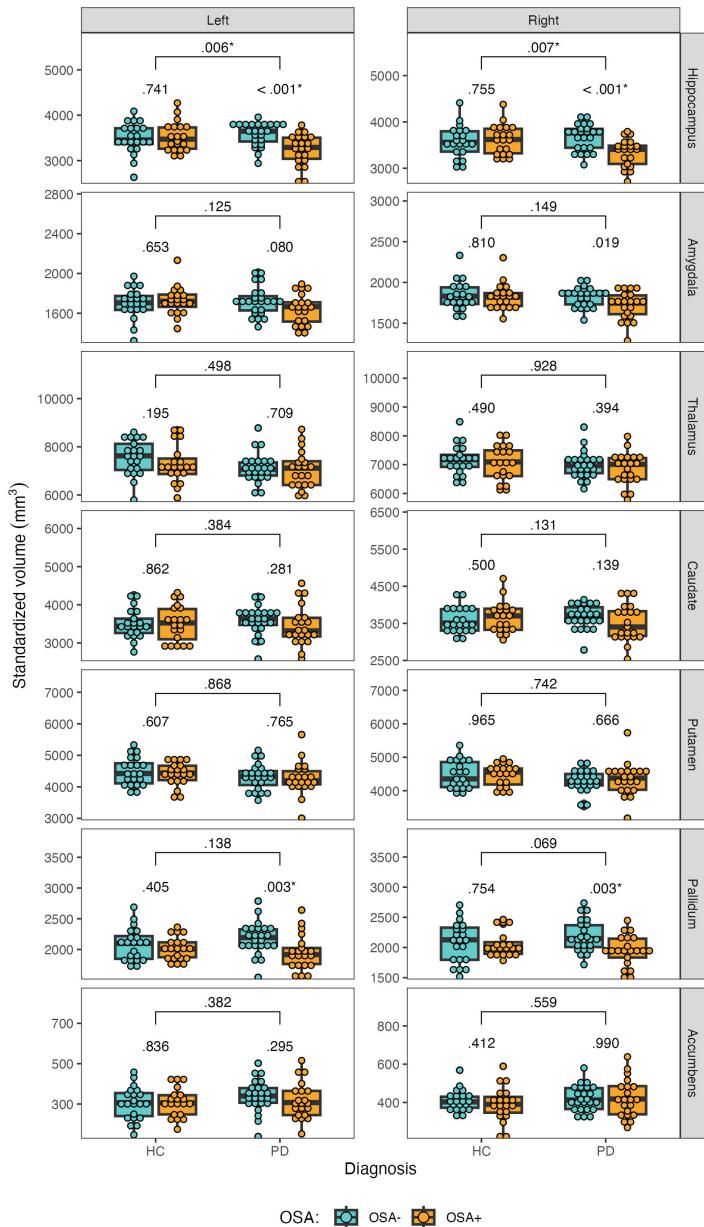
	Left hemisphere							Right hemisphere						
	$\beta^I$	SE	95% CI	t value	p-value	s-value	sig.	$\beta^I$	SE	95% CI	t value	p-value	s-value	sig.
<b>Accumbens</b>														
Group	0.14	0.20	[-0.26, 0.54]	0.712	.479	1.06		0.10	0.19	[-0.29, 0.48]	0.500	.618	0.69	
OSA	0.10	0.19	[-0.28, 0.49]	0.532	.596	0.75		0.05	0.19	[-0.32, 0.43]	0.286	.775	0.37	
Group * OSA	0.22	0.37	[-0.52, 0.96]	0.594	.554	0.85		-0.24	0.36	[-0.96, 0.47]	-0.679	.499	1.00	

<sup>I</sup>Values based on in-sample standardised outcome predicted by Group, OSA, Group \* OSA interaction, age, sex, and TIV via QR decomposition solution to the least squares problem as implemented in R lm() function. Negative values imply smaller adjusted volume of a subcortical structure in PD compared to HC (Group rows), OSA+ compared to OSA- (OSA row) or smaller OSA+ - OSA- difference in HC compared to PD (Group \* OSA rows), the reverse is true for positive values.

$\beta$ , regression coefficient estimate; SE, standard error; CI, confidence interval; OSA, obstructive sleep apnea; sig., coefficients statistically significantly different from 0 after adjusting the p-value threshold for 5% False Discovery Rate via Benjamini-Hochberg adjustment applied to all statistical tests presented in this table are denoted by asterisk "\*\*".

## Cognitive variables

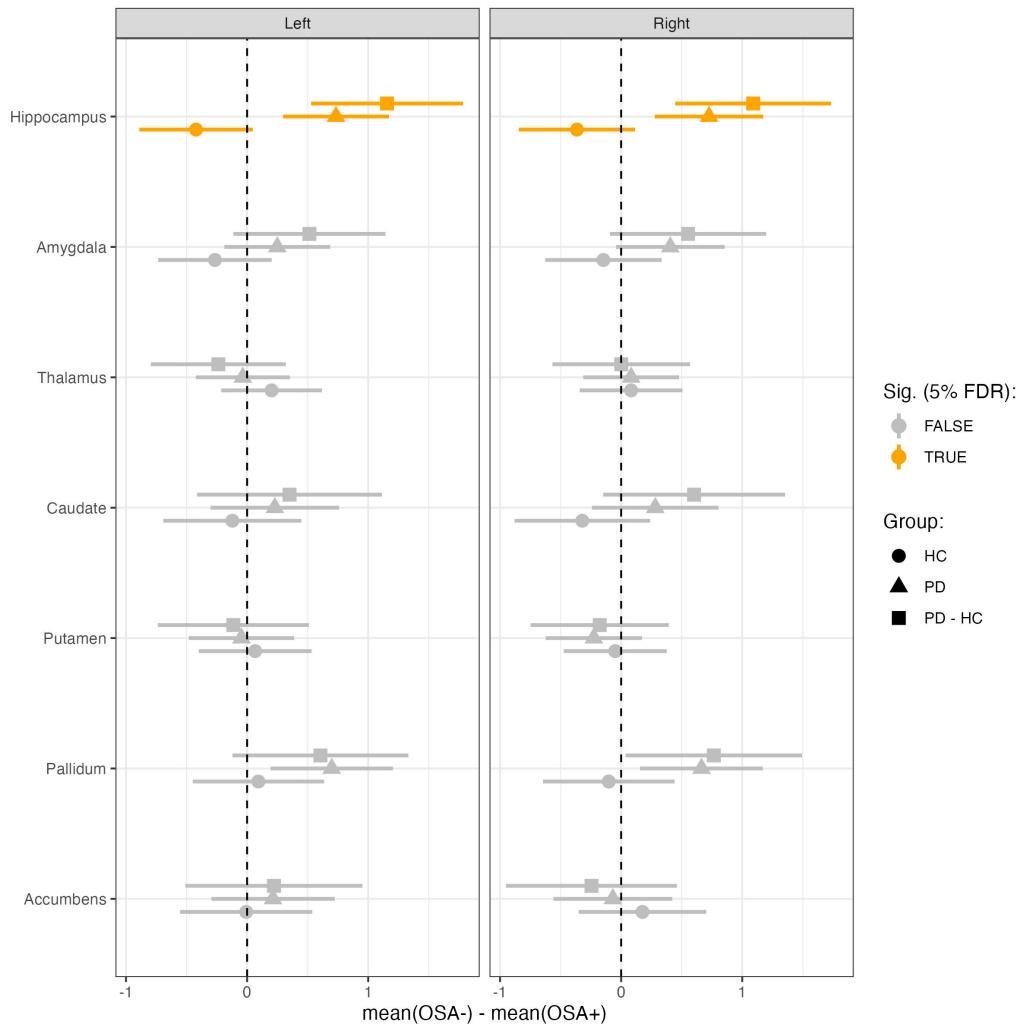
Models of RAVLT-FP (False Positives), PST-W, TMT-B, and both GPT showed deviation from homoscedasticity, and models of RAVLT-FP, RAVLT-FN (False Negatives), TMT-A, PST-D, PST-C and both GPT showed deviation from normality of residuals (see Supplementary Table *cognition\_regression\_coefficients.csv*). None of the diagnosis group/OSA interaction coefficients reach the significance level at the adjusted 5% FDR (see Figure A1 and Table A1). On the other hand, the main effect of diagnostic group was detected using the 5% FDR adjusted threshold in TMT-B and GPT bilaterally (see Figure A2 and Figure A3) with PD patients performing worse on average than HC after adjusting for age, sex and education. Consequently, models of TMT-B and GPT were re-fitted via the distributional regression to check robustness to unequal variance in the data. The distributional models converged successfully and improved upon modelling of data distribution generally (Figure A4) and the group-specific standard deviation in particular (Figure A5). The estimates of model parameters were similar across frequentist and Bayesian distributional models (Table A2) implying these results are likely robust to the unequal variances between HC and PD groups.



**Figure 1**

Raw data (points) as well as median (bar), first and third quartiles (hinges) and 1.5-times interquartile range (whiskers) of standardised volume of subcortical structures being compared between diagnostic and OSA groups. Standardisation for this figure was calculated as (raw volume/TIV)in-sample-mean(TIV). The numbers represent p-values associated with the difference between OSA- and OSA+ conditional on diagnosis group (bottom number) and group/OSA interaction (top number).

Differences with p-value below 5% false discovery rate threshold applied to the set of comparisons presented in this figure are marked by asterisk.



**Figure 2**

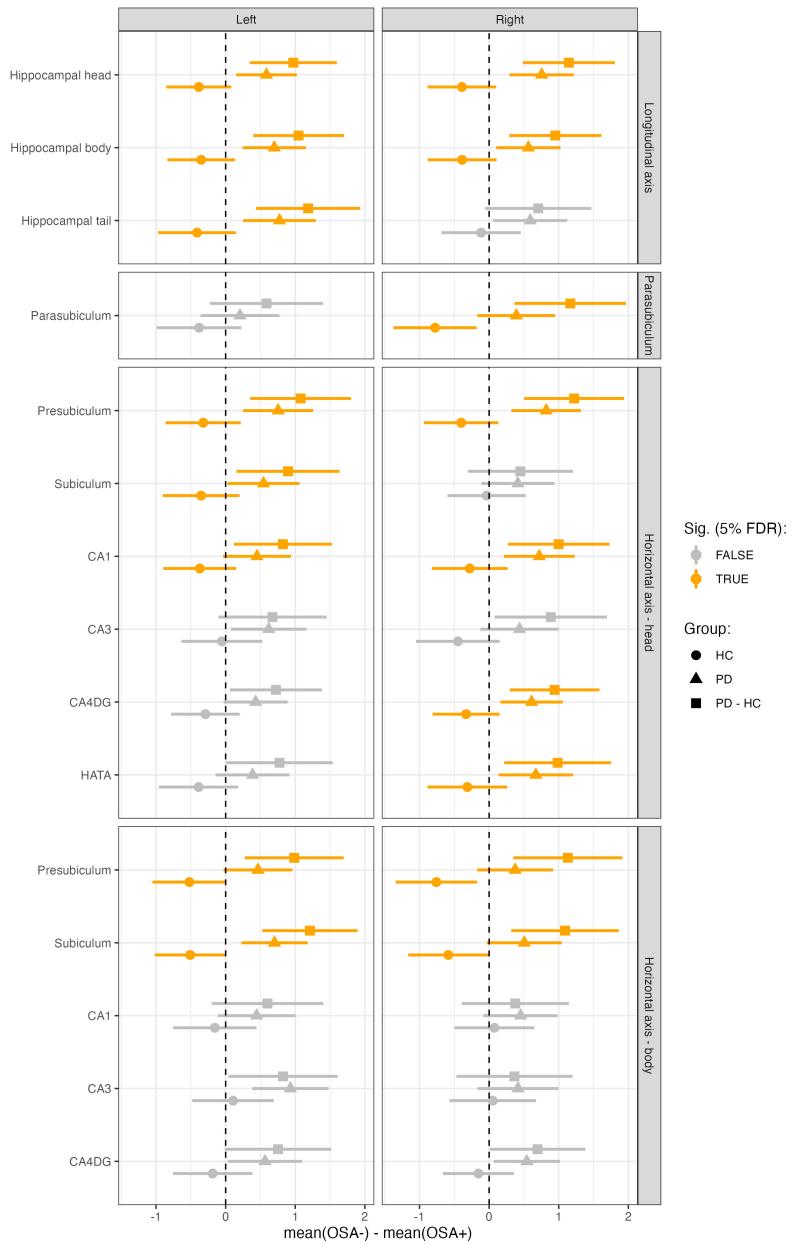
Forest plot showing comparisons of estimated mean age, sex, and TIV adjusted volumes of subcortical structures conditional on group diagnosis and OSA. X-axis represents in-sample standardised difference between mean of OSA- and OSA+ groups estimated from linear regression separately in HC group (circles), PD group (triangles), and difference-in-differences (i.e., interaction) between PD and HC groups (square). Horizontal bars represent non-adjusted 95% confidence intervals. Cases with interaction statistically significantly different from zero after adjusting for 5% false discovery rate via the Benjamini-Hochberg procedure are printed in orange. The figure is based on Supplementary table subcortical\_marginals.csv

**Table 3. Results of regression analyses estimating means of hippocampal fields volume conditional on diagnosis group and obstructive sleep apnea**

Structure	Right hemisphere							Left hemisphere						
	$\beta^1$	SE	95% CI	t value	p value	s value	sig.	$\beta^1$	SE	95% CI	t value	p value	s value	sig.
<b>Longitudinal axis</b>														
Hippocampal head	1.15	0.34	[0.47, 1.82]	3.396	.001	9.86	*	0.97	0.32	[0.34, 1.61]	3.045	.003	8.30	*
Hippocampal body	0.95	0.34	[0.28, 1.62]	2.812	.006	7.33	*	1.05	0.33	[0.39, 1.71]	3.151	.002	8.76	*
Hippocampal tail	0.71	0.39	[-0.07, 1.48]	1.812	.074	3.76		1.18	0.38	[0.42, 1.95]	3.098	.003	8.53	*
<b>Parasubiculum</b>														
Parasubiculum	1.17	0.41	[0.35, 1.98]	2.852	.006	7.49	*	0.59	0.41	[-0.24, 1.41]	1.415	.161	2.64	
<b>Horizontal axis - head</b>														
HATA	0.98	0.39	[0.20, 1.76]	2.514	.014	6.16	*	0.77	0.39	[-0.00, 1.55]	1.988	.050	4.31	
Presubiculum	1.22	0.37	[0.49, 1.95]	3.331	.001	9.57	*	1.08	0.37	[0.34, 1.81]	2.918	.005	7.77	*
Subiculum	0.45	0.38	[-0.32, 1.21]	1.166	.247	2.02		0.90	0.38	[0.14, 1.65]	2.372	.020	5.64	*
CA1	1.00	0.37	[0.26, 1.74]	2.684	.009	6.82	*	0.82	0.36	[0.11, 1.54]	2.296	.024	5.36	*
CA3	0.89	0.41	[0.07, 1.71]	2.153	.034	4.86		0.67	0.40	[-0.12, 1.46]	1.696	.094	3.41	
CA4DG	0.94	0.33	[0.29, 1.60]	2.864	.005	7.55	*	0.72	0.34	[0.05, 1.39]	2.144	.035	4.83	
<b>Horizontal axis - body</b>														
Presubiculum	1.13	0.40	[0.34, 1.93]	2.831	.006	7.41	*	0.98	0.36	[0.26, 1.71]	2.716	.008	6.94	*
Subiculum	1.09	0.39	[0.30, 1.88]	2.763	.007	7.13	*	1.21	0.35	[0.52, 1.90]	3.475	<.001	10.23	*
CA1	0.38	0.39	[-0.41, 1.16]	0.958	.341	1.55		0.60	0.41	[-0.21, 1.42]	1.470	.146	2.78	
CA3	0.36	0.43	[-0.48, 1.21]	0.855	.395	1.34		0.82	0.40	[0.03, 1.62]	2.058	.043	4.54	
CA4DG	0.70	0.35	[0.00, 1.39]	1.997	.049	4.34		0.75	0.39	[-0.02, 1.53]	1.927	.058	4.12	

<sup>1</sup>Values based on in-sample standardised outcome predicted by Group, OSA, Group \* OSA interaction, age, sex, and TIV via QR decomposition solution to the least squares problem as implemented in R lm() function. Negative values imply smaller OSA+ - OSA- difference in HC compared to PD, the reverse is true for positive values.

$\beta$ , regression coefficient estimate; SE, standard error; CI, confidence interval; OSA, obstructive sleep apnea; sig., coefficients statistically significantly different from 0 after adjusting the p-value threshold for 5% False Discovery Rate via Benjamini-Hochberg adjustment applied to all statistical tests presented in this table are denoted by asterisk \*\*.



**Figure 3**

*Forest plot showing comparisons of estimated mean age, sex, and TIV adjusted volumes of hippocampal fields conditional on group diagnosis and OSA. X-axis represents in-sample standardised difference between mean of OSA- and OSA+ groups estimated from linear regression separately in HC group (circles), PD group (triangles), and difference-in-differences (i.e., interaction) between PD and HC groups (square). Horizontal bars represent non-adjusted 95% confidence intervals. Cases with interaction statistically significantly different from zero after adjusting for 5% false discovery rate via the Benjamini-Hochberg procedure are printed in orange. The figure is based on Supplementary table hippo\_base\_marginal\_effects.csv*

## Appendix

**Table A1.** Results of regression analyses estimating means of cognitive performance conditional on diagnosis group and obstructive sleep apnea

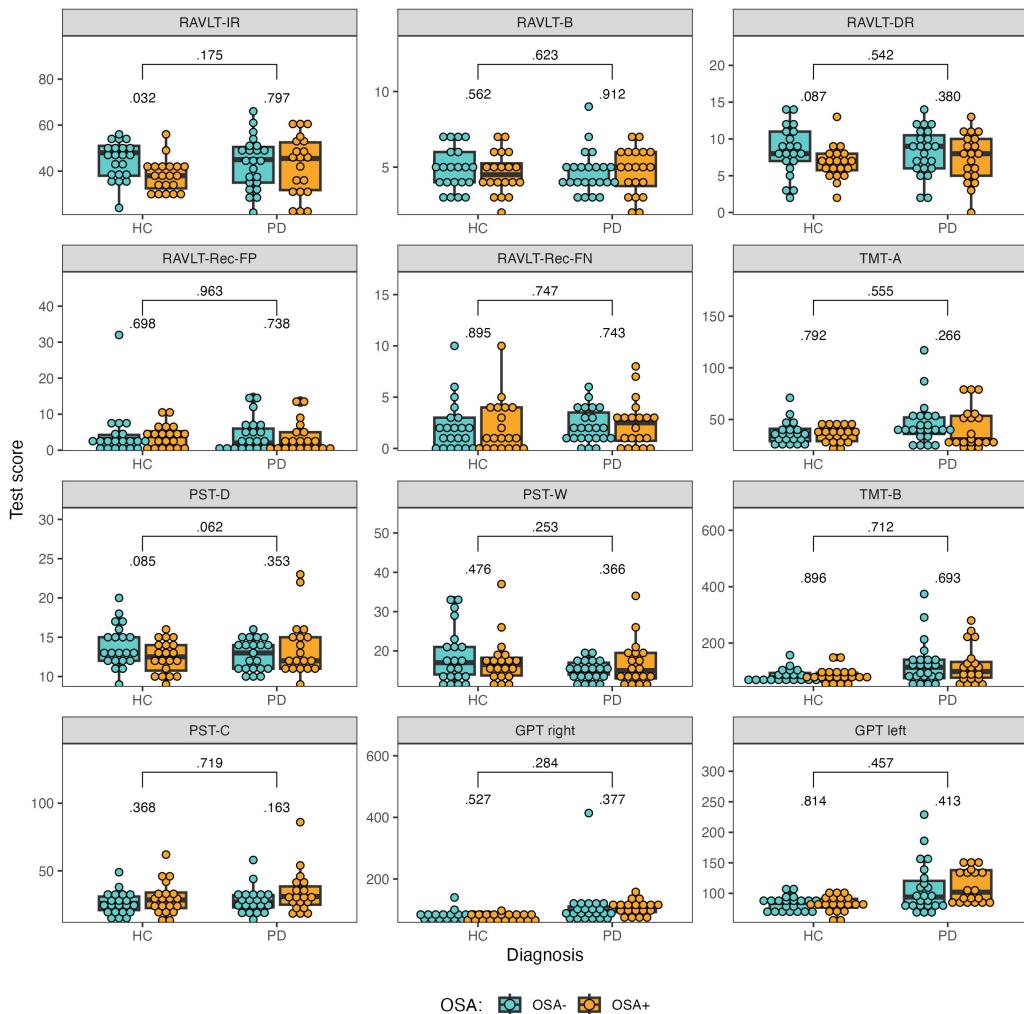
Coefficient	$\beta$	SE	95% CI	t value	p value	s value	sig.
RAVLT-IR							
Group	0.10	0.18	[-0.25, 0.46]	0.590	.557	0.84	
OSA	0.01	0.18	[-0.35, 0.38]	0.080	.937	0.09	
Group * OSA	-0.45	0.35	[-1.15, 0.26]	-1.268	.208	2.26	
RAVLT-B							
Group	-0.02	0.21	[-0.43, 0.40]	-0.091	.928	0.11	
OSA	-0.11	0.22	[-0.54, 0.33]	-0.489	.626	0.67	
Group * OSA	-0.23	0.42	[-1.06, 0.60]	-0.560	.577	0.79	
RAVLT-DR							
Group	0.04	0.20	[-0.36, 0.44]	0.181	.857	0.22	
OSA	0.14	0.21	[-0.28, 0.55]	0.658	.512	0.97	
Group * OSA	-0.16	0.40	[-0.95, 0.64]	-0.388	.699	0.52	
RAVLT-Rec-FP							
Group	-0.07	0.22	[-0.50, 0.36]	-0.325	.746	0.42	
OSA	0.28	0.23	[-0.17, 0.73]	1.242	.218	2.20	
Group * OSA	-0.08	0.43	[-0.95, 0.79]	-0.182	.856	0.22	
RAVLT-Rec-FN							
Group	0.12	0.22	[-0.31, 0.56]	0.566	.573	0.80	
OSA	0.14	0.23	[-0.31, 0.60]	0.627	.533	0.91	
Group * OSA	-0.23	0.44	[-1.10, 0.65]	-0.511	.611	0.71	
TMT-A							
Group	-0.34	0.21	[-0.76, 0.08]	-1.598	.114	3.13	

**Table A1.** Results of regression analyses estimating means of cognitive performance conditional on diagnosis group and obstructive sleep apnea

**Table A1.** Results of regression analyses estimating means of cognitive performance conditional on diagnosis group and obstructive sleep apnea

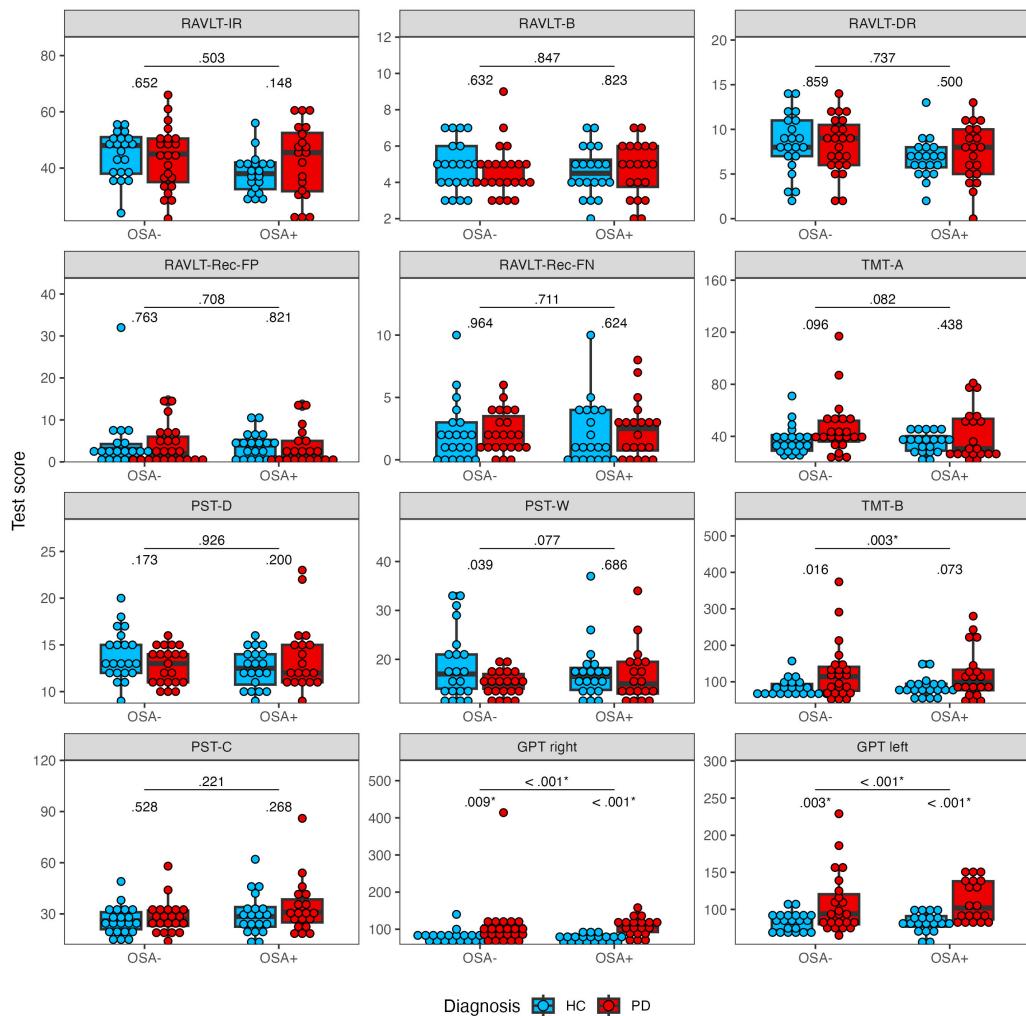
Coefficient	$\beta$	SE	95% CI	t value	p value	s value	sig.
Group	-0.93	0.19	[-1.30, -0.56]	-5.014	< .001	18.17	*
OSA	-0.12	0.19	[-0.51, 0.26]	-0.639	.525	0.93	
Group * OSA	0.28	0.37	[-0.46, 1.02]	0.756	.452	1.15	

$\beta$ , regression coefficient estimate; SE, standard error; CI, confidence interval; OSA, obstructive sleep apnea; sig., coefficients statistically significantly different from 0 after adjusting the p-value threshold for 5% False Discovery Rate via Benjamini-Hochberg adjustment applied to all statistical tests presented in this table are denoted by asterisk \*\*.



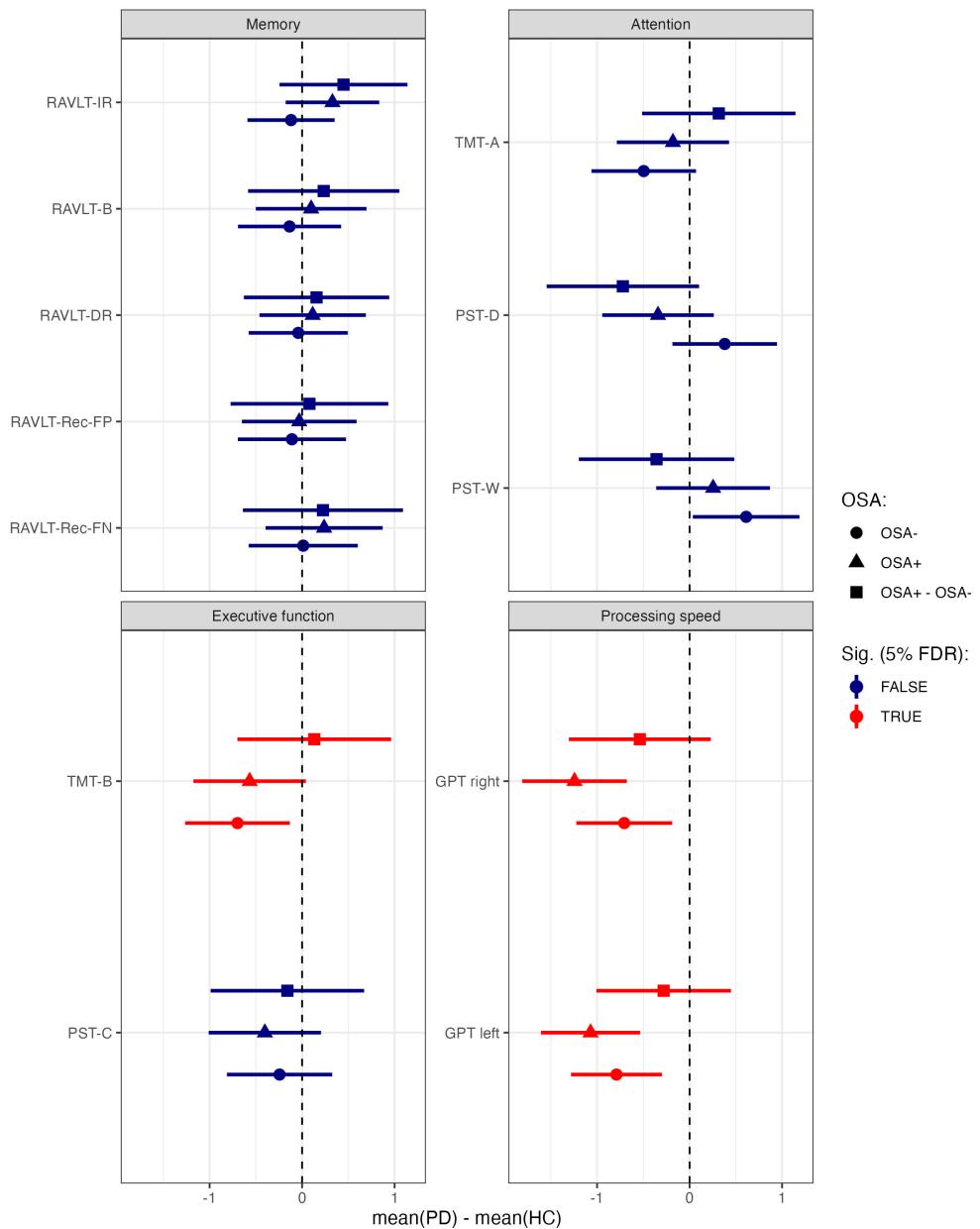
**Figure A1**

Raw data (points) as well as median (bar), first and third quartiles (hinges) and 1.5-times interquartile range (whiskers) of indexes of cognitive functions being compared between diagnostic and OSA groups. The numbers represent p-values associated with the difference between OSA- and OSA+ conditional on diagnosis group (bottom number) and group/OSA interaction (top number). Non of the p-value were 5% false discovery rate threshold applied to the set of comparisons presented in this figure.



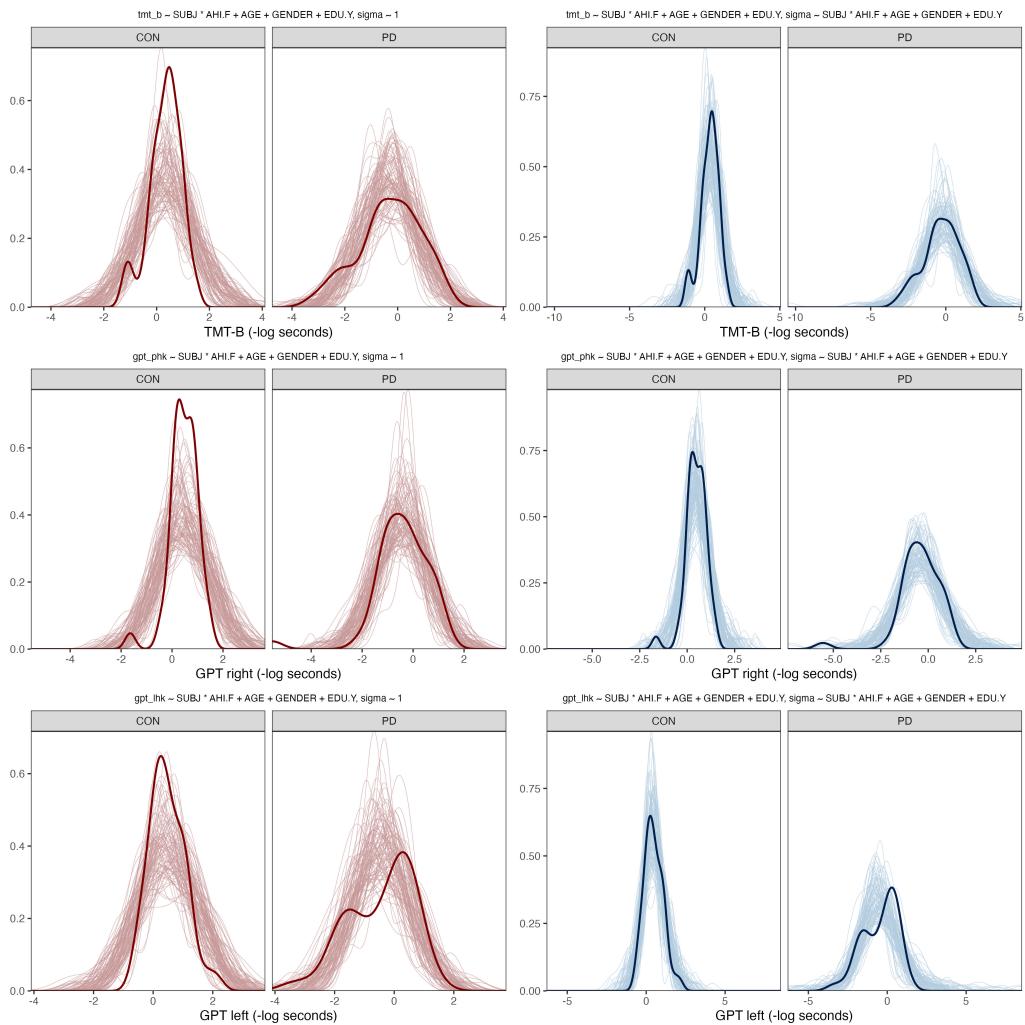
**Figure A2**

Alternative presentation of raw data (points), median (bar), first and third quartiles (hinges) and 1.5-times interquartile range (whiskers) of indexes of cognitive functions being compared between diagnostic and OSA groups. The numbers represent p-values associated with the difference between PD and HC conditional on OSA (bottom number) and the overall difference between PD and HC across OSA groups (top number above brackets). Differences with p-value below 5% false discovery rate threshold applied to the set of comparisons presented in this figure are marked by asterisk.



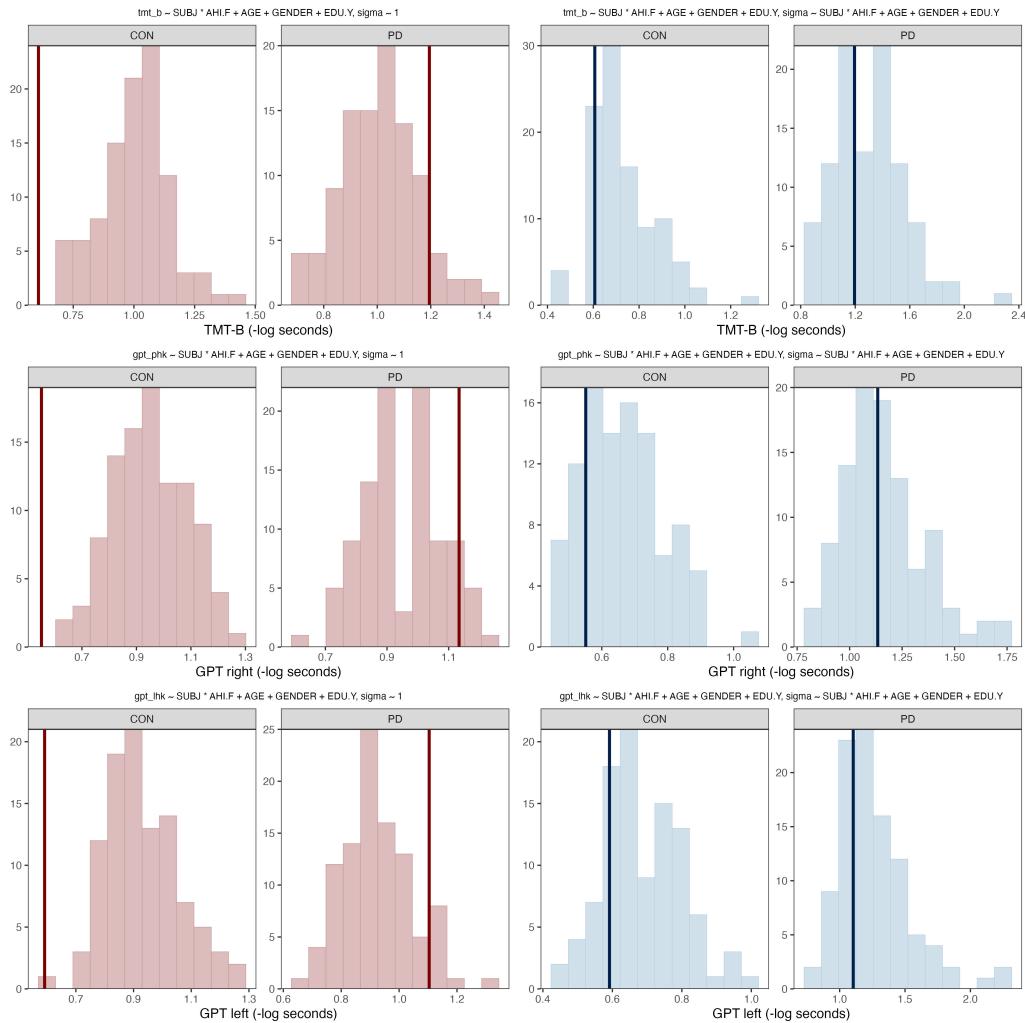
**Figure A3**

Forest plot showing comparisons of estimated mean age, sex, and education adjusted cognitive performance conditional on group diagnosis and OSA. X-axis represents in-sample standardised difference between mean of PD and HC groups estimated from linear regression separately in OSA- group (circles), OSA+ group (triangles), and difference-in-differences (i.e., interaction) between OSA- and OSA+ groups (square). Horizontal bars represent non-adjusted 95% confidence intervals. Cases with main effect of diagnosis group statistically significantly different from zero after adjusting for 5% false discovery rate via the Benjamini-Hochberg procedure are printed in red. Note that the colour scheme is different from previous figures because the current figure shows different contrasts. The figure is based on Supplementary table cognition\_marginals.csv



**Figure A4**

Posterior predictive check of data distribution predicted by classical (red) and distributional (blue) models of response times variables. Thick lines represent observed data distribution, thin lines represent one hundred random posterior predictive distributions of the model specified above each panel.



**Figure A5**

Posterior predictive check of standard deviation (SD) predicted by classical (red) and distributional (blue) models of response times variables. Thick lines represent SD in observed data, whereas histograms represent SD from all posterior predictive distributions of the model specified above each panel.

**Table A2.** Comparison of the frequentist regression coefficients and Bayesian distributional regression coefficients accounting for heteroscedasticity

Coefficient	Bayesian						Frequentist		
	Classical			Distributional					
	$\beta^1$	SE	95% PPI	$\beta^1$	SE	95% PPI	$\beta^1$	SE	95% CI
<b>TMT-B</b>									
Group	-0.61	0.21	[-1.02, -0.19]	-0.57	0.21	[-0.99, -0.16]	-0.63	0.21	[-1.05, -0.21]
OSA	-0.15	0.22	[-0.58, 0.28]	-0.12	0.21	[-0.52, 0.28]	-0.16	0.22	[-0.60, 0.28]
Group * OSA	-0.12	0.40	[-0.89, 0.64]	-0.22	0.40	[-0.98, 0.57]	-0.13	0.42	[-0.98, 0.71]
<b>GPT right</b>									
Group	-0.94	0.20	[-1.30, -0.54]	-0.87	0.17	[-1.20, -0.53]	-0.98	0.20	[-1.37, -0.59]
OSA	-0.12	0.20	[-0.52, 0.26]	-0.03	0.17	[-0.37, 0.29]	-0.14	0.20	[-0.54, 0.27]
Group * OSA	0.45	0.36	[-0.24, 1.14]	0.38	0.32	[-0.26, 1.00]	0.54	0.39	[-0.24, 1.32]
<b>GPT left</b>									
Group	-0.89	0.19	[-1.26, -0.54]	-0.94	0.18	[-1.28, -0.59]	-0.93	0.19	[-1.30, -0.56]
OSA	-0.11	0.19	[-0.49, 0.25]	-0.09	0.18	[-0.43, 0.28]	-0.12	0.19	[-0.51, 0.26]
Group * OSA	0.24	0.34	[-0.43, 0.89]	0.22	0.34	[-0.47, 0.88]	0.28	0.37	[-0.46, 1.02]

<sup>1</sup>Values based on in-sample standardised outcome predicted by Group, OSA, Group \* OSA interaction, age, sex, and TIV via QR decomposition solution to the least squares problem as implemented in R lm() function. Negative values imply smaller OSA+ - OSA- difference in HC compared to PD, the reverse is true for positive values.

$\beta$ , regression coefficient estimate; SE, standard error; CI, confidence interval; PPI, equal-tailed posterior probability interval; OSA, obstructive sleep apnea.

## References

- Arel-Bundock, Vincent, Noah Greifer, and Andrew Heiss. Forthcoming. "How to Interpret Statistical Models Using *marginaleffects* in R and Python." *Journal of Statistical Software*, Forthcoming.
- Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1): 289–300. <https://doi.org/https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Bezdicek, Ondrej, Jiri Lukavsky, Hana Stepankova, Tomas Nikolai, Bradley N. Axelrod, Jiri Michalec, Evžen Růžička, and Miloslav Kopecek. 2015. "The Prague Stroop Test: Normative Standards in Older Czech Adults and Discriminative Validity for Mild Cognitive Impairment in Parkinson's Disease." *Journal of Clinical and Experimental Neuropsychology* 37 (8): 794–807. <https://doi.org/10.1080/13803395.2015.1057106>.
- Bezdicek, Ondrej, L. Motak, B. N. Axelrod, M. Preiss, T. Nikolai, M. Vyhnalek, A. Poreh, and E. Ruzicka. 2012. "Czech Version of the Trail Making Test: Normative Data and Clinical Utility." *Archives of Clinical Neuropsychology* 27 (8): 906–14. <https://doi.org/10.1093/arclin/acs084>.
- Bezdicek, Ondrej, Hana Stepankova, Bradley N. Axelrod, Tomas Nikolai, Zdenek Sulc, Robert Jech, Evžen Růžička, and Miloslav Kopecek. 2017. "Clinimetric Validity of the Trail Making Test Czech Version in Parkinson's Disease and Normative Data for Older Adults." *The Clinical Neuropsychologist* 31 (sup1): 42–60. <https://doi.org/10.1080/13854046.2017.1324045>.
- Bezdicek, Ondrej, Hana Stepankova, Ladislav Moták, Bradley N. Axelrod, John L. Woodard, Marek Preiss, Tomáš Nikolai, Evžen Růžička, and Amir Poreh. 2014. "Czech Version of Rey Auditory Verbal Learning Test: Normative Data." *Aging, Neuropsychology, and Cognition* 21 (6): 693–721. <https://doi.org/10.1080/13825585.2013.865699>.
- Bezdicek, Ondrej, Zdenek Sulc, Tomas Nikolai, Hana Stepankova, Miloslav Kopecek, Robert Jech, and Evžen Růžička. 2017. "A Parsimonious Scoring and Normative Calculator for the Parkinson's Disease Mild Cognitive Impairment Battery." *The Clinical Neuropsychologist* 31 (6-7): 1231–47. <https://doi.org/10.1080/13854046.2017.1293161>.
- Bürkner, Paul-Christian. 2017. "Brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80 (1). <https://doi.org/10.18637/jss.v080.i01>.
- Cole, Stephen R, Jessie K Edwards, and Sander Greenland. 2021. "Surprise!" *American Journal of Epidemiology* 190 (2): 191–93.

- Frydrychová, Zuzana, Miloslav Kopeček, Ondrej Bezdicek, and Hana Georgi Stepankova. 2018. "Czech normative study of the Revised Rey Auditory Verbal Learning Test (RAVLT) in older adults." *Ceskoslovenska Psychologie* 62 (4): 330–49.
- Greenland, Sander. 2019. "Valid p-Values Behave Exactly as They Should: Some Misleading Criticisms of p-Values and Their Resolution with s-Values." *The American Statistician* 73 (sup1): 106–14.
- Kløve, Hallgrim. 1963. "Clinical Neuropsychology." *Medical Clinics of North America* 47 (6): 1647–58. [https://doi.org/10.1016/S0025-7125\(16\)33515-5](https://doi.org/10.1016/S0025-7125(16)33515-5).
- Kopecek, Miloslav, Hana Stepankova, Jiri Lukavsky, Daniela Ripova, Tomas Nikolai, and Ondrej Bezdicek. 2017. "Montreal Cognitive Assessment (MoCA): Normative Data for Old and Very Old Czech Adults." *Applied Neuropsychology: Adult* 24 (1): 23–29. <https://doi.org/10.1080/23279095.2015.1065261>.
- Litvan, Irene, Jennifer G. Goldman, Alexander I. Tröster, Ben A. Schmand, Daniel Weintraub, Ronald C. Petersen, Brit Mollenhauer, et al. 2012. "Diagnostic Criteria for Mild Cognitive Impairment in Parkinson's Disease: Movement Disorder Society Task Force Guidelines." *Movement Disorders* 27 (3): 349–56. <https://doi.org/10.1002/mds.24893>.
- Nasreddine, Ziad S., Natalie A. Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L. Cummings, and Howard Chertkow. 2005. "The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool for Mild Cognitive Impairment." *Journal of the American Geriatrics Society* 53 (4): 695–99. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>.
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Stan Development Team. 2020. "Stan Modeling Language Users Guide and Reference Manual, Version 2.21.0." <http://mc-stan.org/>.
- Wagenmakers, Eric-Jan, and Scott Brown. 2007. "On the Linear Relation Between the Mean and the Standard Deviation of a Response Time Distribution." *Psychological Review* 114 (3): 830. <https://doi.org/10.1037/0033-295X.114.3.830>.
- Wasserstein, Ronald L, and Nicole A Lazar. 2016. "The ASA Statement on p-Values: Context, Process, and Purpose." *The American Statistician*. Taylor & Francis.