

Numerical vs. Anatomical Variability: Impact on Numerical Reliability on MRI measures of Parkinson’s Disease

Yohan Chatelain, Andrzej Sokołowski, Madeleine Sharp, Jean-Baptiste Poline, Tristan Glatard

July 2, 2025

Abstract

Reproducibility in neuroimaging is critically hampered by computational variability within analysis software. Here, we quantify the numerical instability of FreeSurfer, a ubiquitous tool, using structural MRI data from the Parkinson’s Progression Markers Initiative. By introducing controlled perturbations with Monte Carlo Arithmetic, we assessed the variability of cortical and subcortical measurements. We found that numerical variability in FreeSurfer is substantial, with key structural metrics often precise to only a single significant digit. This computational noise was frequently comparable to or greater than the biological variation between Parkinson’s disease patients and healthy controls. Consequently, the statistical significance of group differences and correlations with clinical severity fluctuated dramatically across identical analyses. Our results demonstrate that subtle computational errors can produce unreliable findings in clinical neuroimaging. Addressing such numerical instability is essential for developing robust biomarkers for neurological disorders like Parkinson’s disease.

1 Introduction

Neuroimaging reproducibility has emerged as a critical challenge in neuroscience research. While inter-software variability is well-documented [BNHC⁺20, GHJ⁺12, BBD⁺21], within-version numerical variability—small output variations from identical software runs—remains underexplored despite potentially significant clinical implications. Numerical variability arises from computational factors including floating-point precision, parallel processing, and random initializations. In Parkinson’s disease (PD) research, where MRI-derived metrics like cortical thickness and subcortical volumes serve as potential biomarkers, such variability could obscure subtle disease-related changes and compromise statistical reliability.

Previous studies have demonstrated substantial between-version differences in FreeSurfer outputs [HPZ⁺23], but the impact of computational uncertainty within single software versions on clinical associations remains unclear. This gap is particularly concerning for PD research, where establishing reliable brain-behavior relationships is essential for developing neuroimaging biomarkers. Despite promising associations between MRI-derived metrics and PD severity, no neuroimaging biomarkers are widely accepted for clinical diagnosis or monitoring. Measurement variability across studies undermines reliability and generalizability, hindering translation to clinical practice. This computational uncertainty could significantly impact PD research by: (1) masking subtle disease-related changes essential for early detection, (2) compromising statistical power for detecting group differences and clinical correlations, and (3) reducing reproducibility across studies using identical analysis pipelines.

Here, we investigate numerical variability in FreeSurfer 7.3.1 using Monte Carlo Arithmetic to simulate realistic computational perturbations. We introduce the Numerical-Anatomical Variability Ratio (NAVR) to quantify computational uncertainty relative to biological variation and derive its theoretical relationship to statistical effect sizes. Using longitudinal data from the Parkinson’s Progression Markers Initiative, we assess how numerical precision affects group comparisons and clinical correlations in PD research. Specifically, our aims include: (1) quantifying how computational uncertainty affects group difference detection between PD patients and healthy controls, (2) assessing numerical variability effects on brain-behavior correlations with clinical measures (UPDRS scores), and (3) developing the NAVR framework to predict statistical reliability from computational precision. Our findings will inform strategies for mitigating numerical variability effects and enhancing reproducibility in clinical neuroimaging studies.

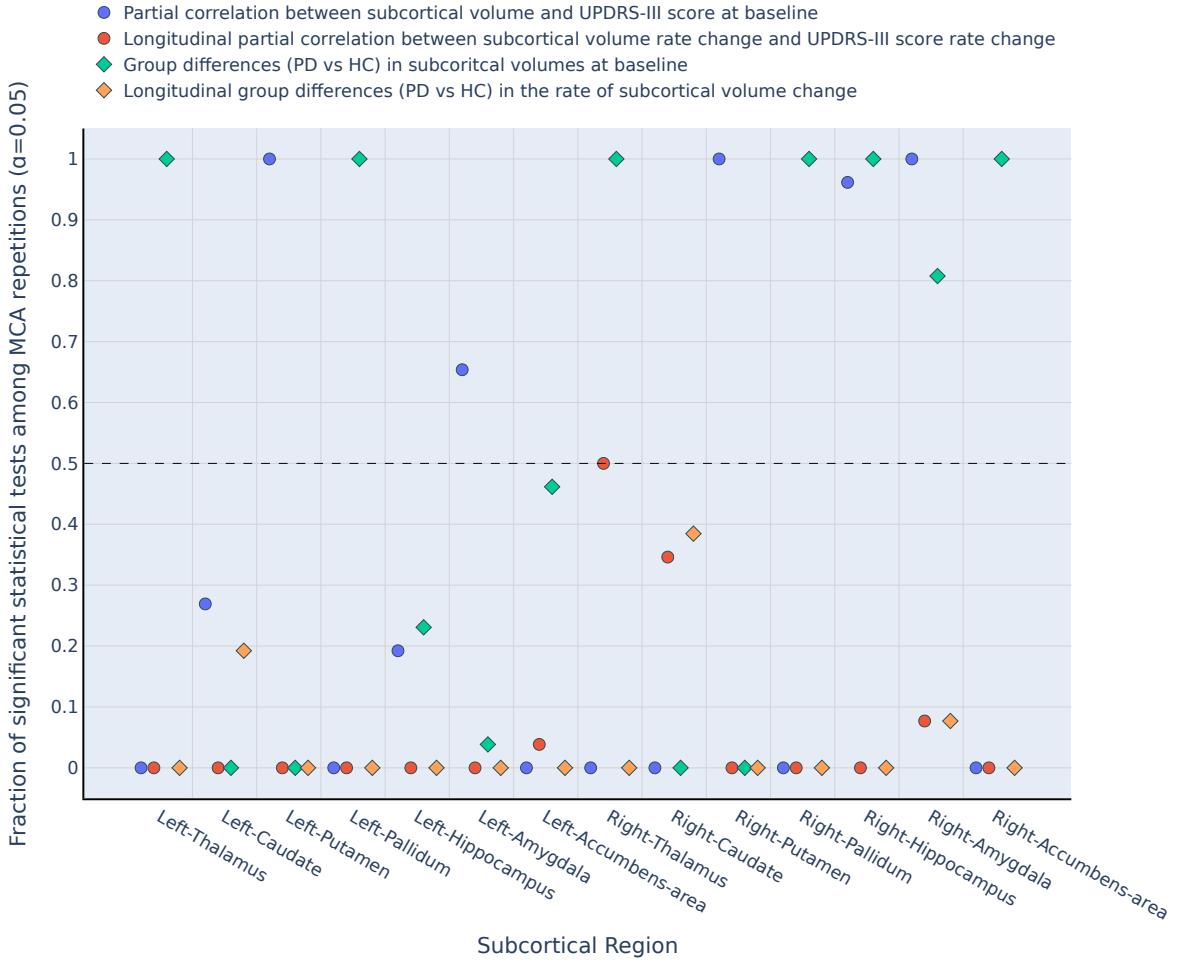


Figure 1: Proportion of statistically significant tests ($p < 0.05$) across the 26 MCA repetitions for subcortical volume measures.

2 Results

PD and HC groups showed no significant age differences ($p > 0.05$) but differed in education ($t = -2.05$, $p = 0.04$) and sex distribution ($\chi^2 = 4.15$, $p = 0.04$). The longitudinal cohort showed no significant demographic differences between groups (Table 2).

2.1 Numerical variability impacts MRI derived findings

Statistical significance proportions across 26 MCA repetitions varied substantially for subcortical volumes (Figure ??). Ratios near 0.5 indicated maximal uncertainty, while values approaching 0 or 1 suggested consistent results across computational variations.

Effect size distributions showed notable variability across MCA repetitions (Figure 3). Partial correlation coefficients and F-statistics from ANCOVA analyses demonstrated spread around standard IEEE-754 results (red markers), indicating that numerical precision affects both statistical significance and effect size estimation.

2.2 NAVR reveals region-specific numerical instabilities

NAVR values varied substantially across brain regions, with some showing computational uncertainty variation (Figure 5). Regions with high NAVR values indicate areas where numerical precision limitations may compromise the detection of true anatomical differences.

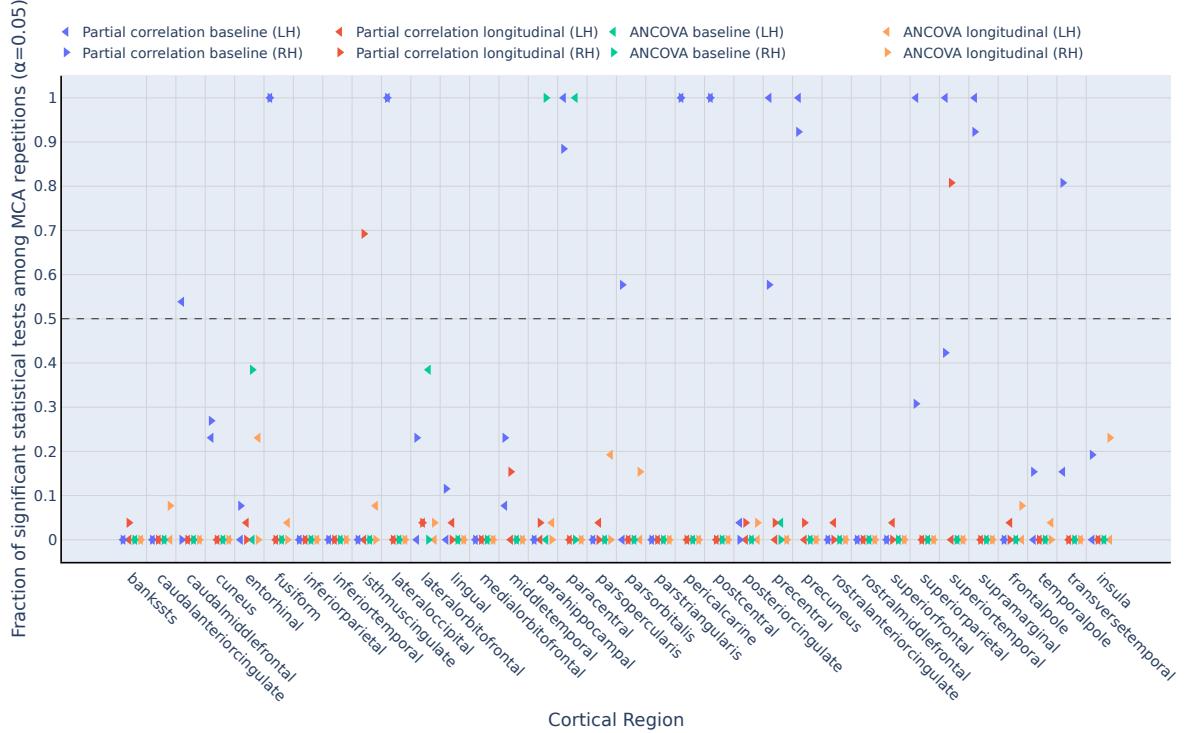


Figure 2: Proportion of statistically significant tests ($p < 0.05$) across the 26 MCA repetitions for cortical thickness measures.

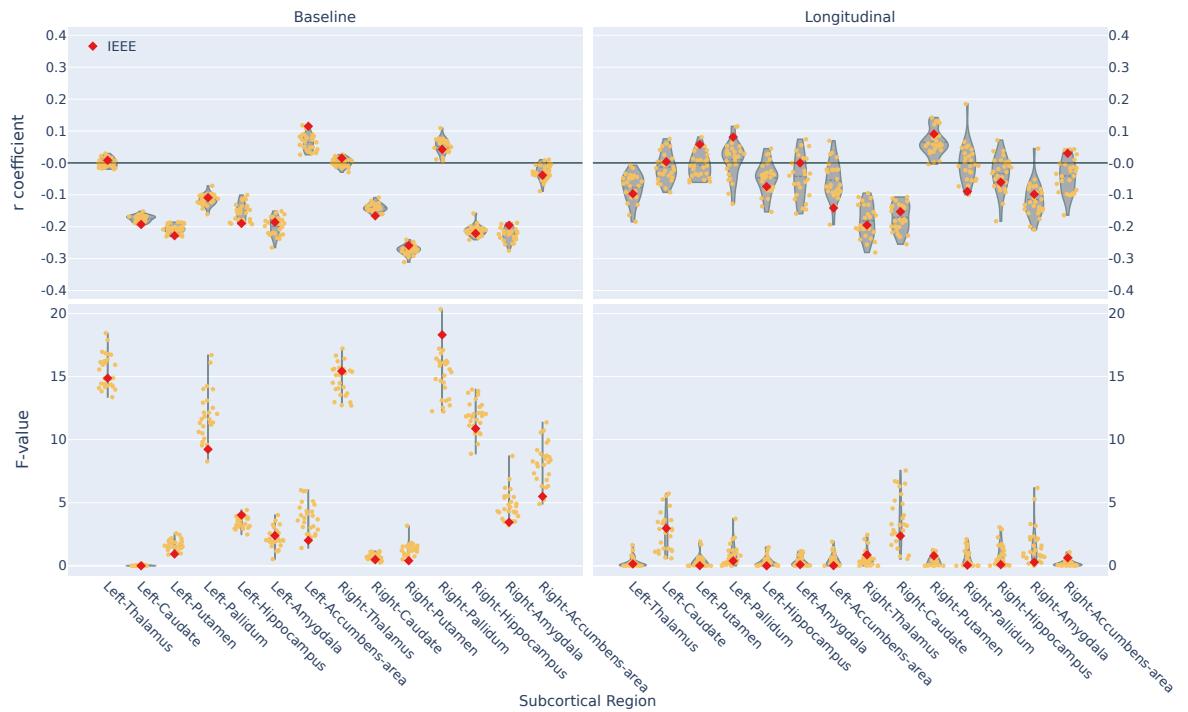
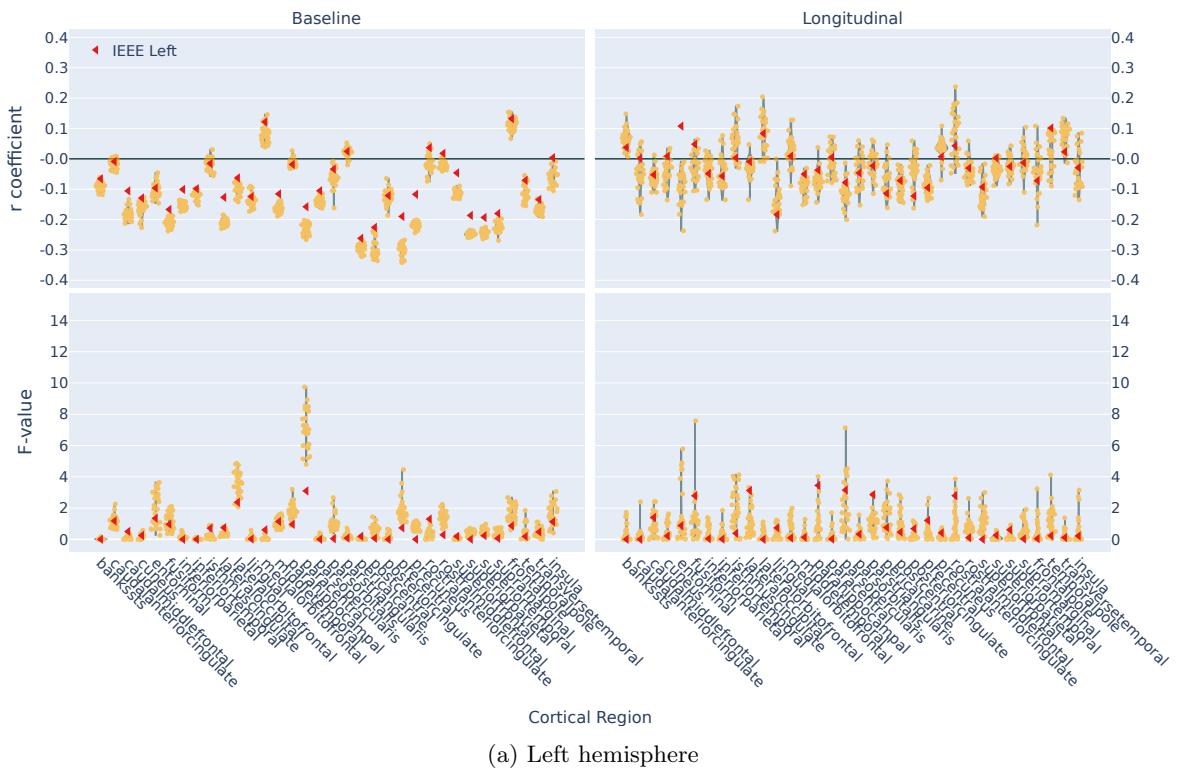
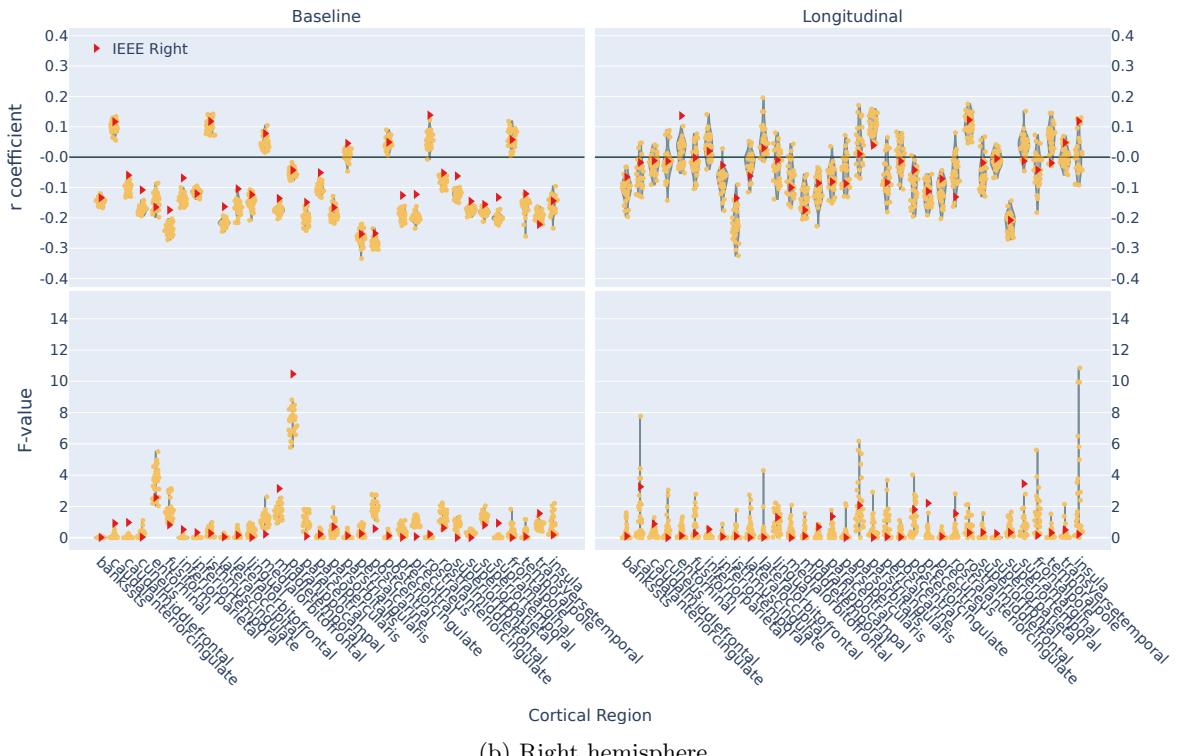


Figure 3: Distribution of partial correlation coefficients (r-values) and F-statistics from ANCOVA across MCA repetitions for subcortical volume measures. Red dots represent the IEEE results. The top row shows r-values, while the bottom row shows F-values. The left column represents baseline analysis, and the right column represents longitudinal analysis.



(a) Left hemisphere



(b) Right hemisphere

Figure 4: Distribution of partial correlation coefficients for cortical thickness across all subjects and regions. Red triangles indicate the IEEE-754 run for reference. The distribution shows the variability in the coefficients, with some regions exhibiting higher consistency than others.

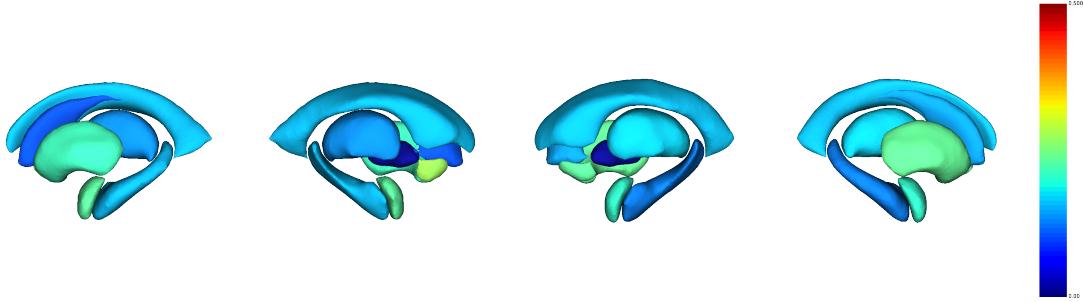


Figure 5: Numerical-Anatomical Variability Ratio (NAVR) for subcortical volumes across regions and groups. Higher NAVR values indicate greater computational uncertainty relative to biological variation.

2.3 Cross-sectional numerical precision analysis

FreeSurfer 7.3.1 showed limited numerical precision across all cortical measures: 1.49 ± 0.27 significant digits for thickness, 1.03 ± 0.27 for surface area, and 1.00 ± 0.28 for volume (Figures 8, 9). These values indicate measurements are typically precise to only one decimal place, with some instances showing complete precision loss.

Regional consistency was observed within each metric type, with thickness showing the highest precision (range: 1.05-1.75 digits) compared to area (0.58-1.35 digits) and volume (0.52-1.36 digits). Dice coefficients revealed substantial inter-subject variability, particularly in vessel regions (Figure 7).

- Presentation of the basic numerical summary of the collected data for both HC and PD groups.

To assess statistical significance, we employed permutation tests. Permutation test provides a non-parametric test to evaluate whether a given measure significantly deviates from the population. Variability introduced by MCA reveals the numerical instability of FreeSurfer 7.3.1 shows that the significance of cortical regions fluctuated across repetitions.

Specifically, only two regions were found to be statistically significant in two repetitions out of 26, suggesting a lack of statistical power. The surface sizes of significant regions varied considerably, with examples such as $1005.95 \pm 39.34 mm^2$ and $996.83 \pm 42.91 mm^2$.

Group Analysis No clusters were found to be statistically significant using the permutation test (1000 permutations).

Correlation Analysis The following regions were identified using a permutation test (1000 permutations) for the correlation analysis. Each region was significant in only one MCA repetition, further demonstrating a lack of replicability.

3 Discussion

Our analysis reveals significant numerical instability in FreeSurfer 7.3.1, with cortical measurements showing limited precision (1-1.5 significant digits) that substantially impacts statistical reliability in neuroimaging studies. These precision limitations pose particular challenges for detecting subtle disease-related changes in conditions like Parkinson’s disease.

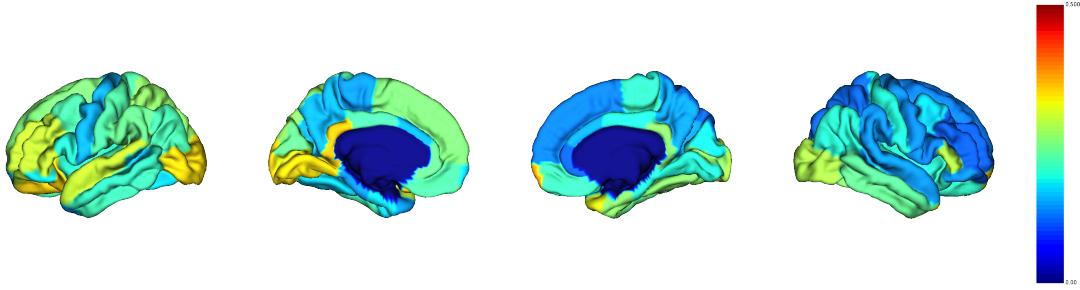


Figure 6: Numerical-Anatomical Variability Ratio (NAVR) for cortical thickness across regions and groups. Higher NAVR values indicate greater computational uncertainty relative to biological variation. The color scale indicates the NAVR value, with warmer colors indicating higher NAVR values.

Table 1: Significant regions identified using a permutation test (1000 permutations) for the correlation analysis.

Region	Size (mm^2)	MNI X	MNI Y	MNI Z	Max	Frequency
L inferior temporal	3337.13	-56.6	-43.6	-18.1	2.8572	1/25
R lingual	1150.84	23.2	-61.3	0.4	4.6403	1/25
R parstriangularis	3265.87	44.4	35.8	3.7	2.8129	1/25

The absence of significant baseline differences between PD and HC groups, combined with inconsistent cluster detection (only 1/26 clusters reproduced across repetitions), demonstrates how numerical variability can compromise reproducibility. The NAVR framework quantifies this relationship, showing that computational uncertainty approaches or exceeds biological variation in many brain regions.

Statistical test consistency varied markedly across MCA repetitions, with methodological choices (Z-test vs. permutation test) further influencing outcome reliability. Effect size distributions showed substantial spread around standard IEEE-754 results, indicating that numerical precision directly affects both significance testing and effect size estimation.

Importantly, inter-subject variability exceeded intra-subject variability, suggesting that FreeSurfer maintains relative consistency across different individuals despite numerical limitations. This supports continued use while highlighting the need for improved computational precision in future neuroimaging software development. This study demonstrates significant numerical limitations in FreeSurfer 7.3.1, with cortical measurements exhibiting only 1-1.5 significant digits of precision. These computational constraints substantially impact statistical reliability and reproducibility in neuroimaging research, particularly for detecting subtle disease-related changes.

Our NAVR framework quantifies the relationship between computational uncertainty and biological variation, revealing that numerical instability approaches or exceeds anatomical variability in many brain regions. This finding has direct implications for statistical power, as demonstrated by inconsistent cluster detection (only 1/26 clusters reproduced) and variable effect sizes across identical analyses.

While inter-subject variability exceeded intra-subject variability—supporting relative consistency across individuals—the absence of significant PD-HC differences and weak clinical correlations highlight how numerical limitations can mask true biological signals. The theoretical relationship between NAVR and Cohen’s d uncertainty provides a framework for predicting statistical reliability based on computational precision.

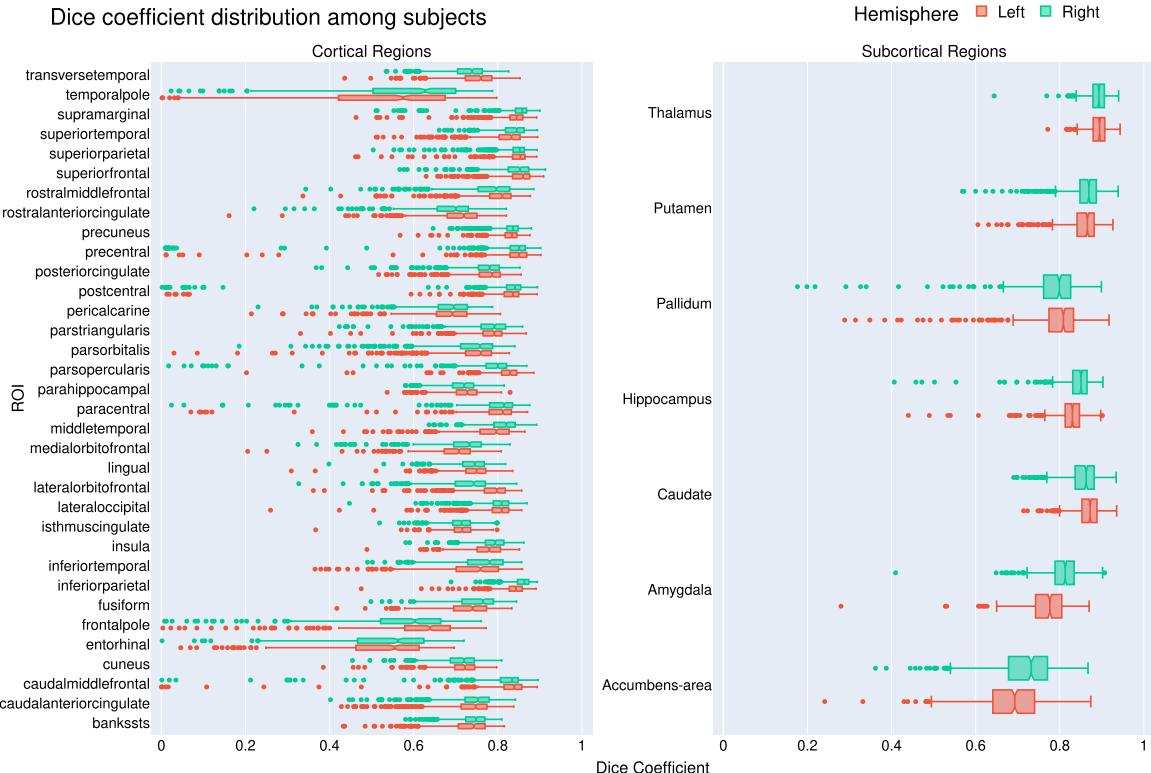


Figure 7: Dice coefficient.

These findings emphasize the critical need for improved numerical precision in neuroimaging software. Future developments should prioritize computational stability to enhance the detection of subtle neurological changes and improve reproducibility across studies. The NAVR framework offers a practical tool for assessing and comparing the numerical reliability of neuroimaging methodologies.

4 Methods

4.1 Numerical variability assessment

We employed Monte Carlo Arithmetic (MCA) [Par97] to quantify numerical instability in FreeSurfer computations. MCA introduces controlled random perturbations into floating-point operations, simulating rounding errors that occur across different computational environments. This stochastic approach enables systematic assessment of result stability by measuring variation across multiple runs of identical analyses.

We used Fuzzy-libm [SCKG21], which extends MCA to mathematical library functions (`exp`, `log`, `sin`, `cos`) through Verificarlo [DdOCP16], an LLVM-based compiler. Virtual precision parameters were set to 53 bits for double precision and 24 bits for single precision to simulate realistic machine-level precision errors.

4.2 Participants

We analyzed data from the Parkinson’s Progression Markers Initiative (PPMI), a multi-site longitudinal study. From 316 initial participants, we selected 125 Parkinson’s disease patients without mild cognitive impairment (PD-non-MCI) and 106 healthy controls (HC) with complete longitudinal T1-weighted MRI data. PD-MCI patients were excluded to avoid confounding effects of cognitive impairment.

Inclusion criteria required: (1) primary PD diagnosis or healthy control status, (2) availability of two visits with T1-weighted scans, and (3) absence of other neurological diagnoses. PD severity was assessed

Significant digits distribution among subjects

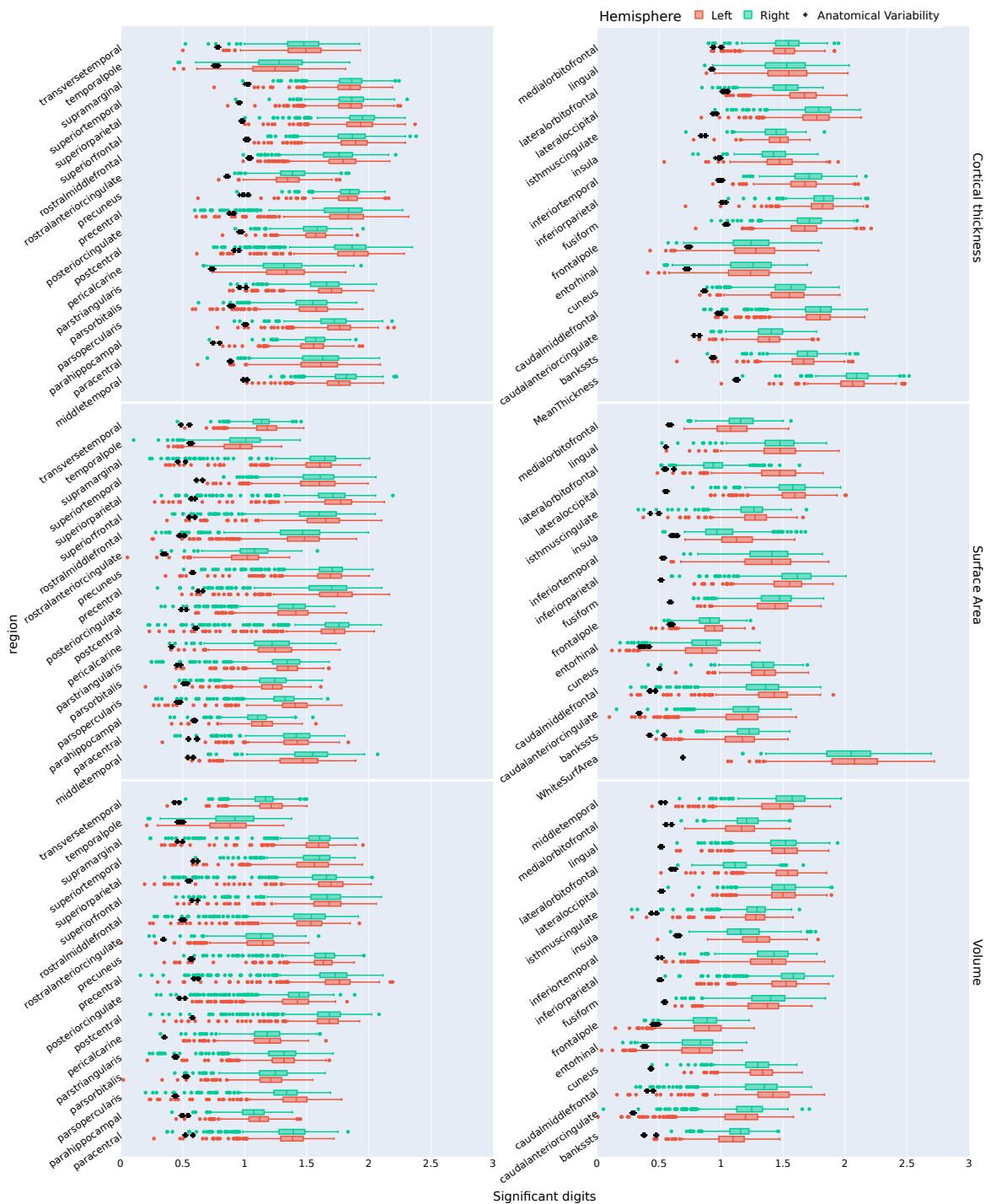


Figure 8: Number of significant digits for each cortical region and metric.

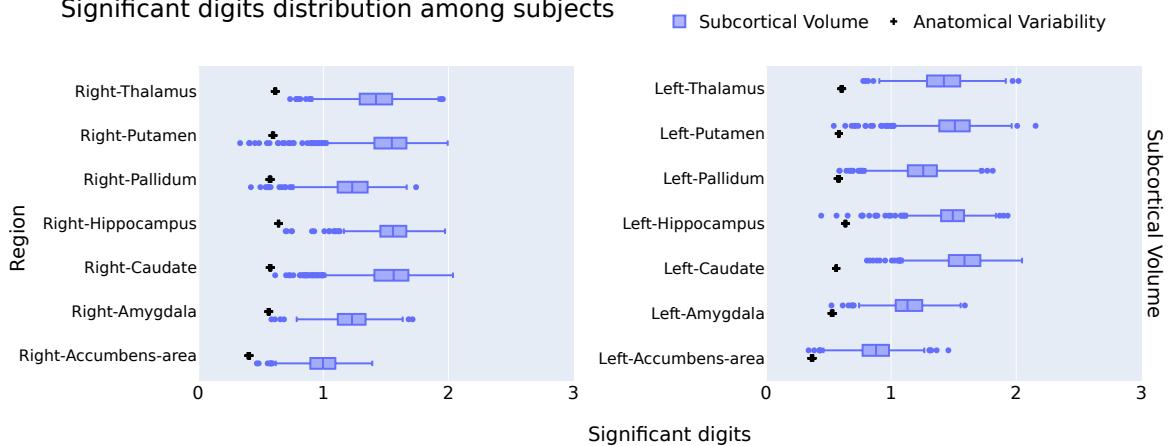


Figure 9: Number of significant digits of subcortical volume for each ROI.

using the Unified Parkinson’s Disease Rating Scale (UPDRS). The study received ethics approval from participating institutions, and all participants provided written informed consent (Table 2).

Cohort	HC	PD-non-MCI
n	103	121
Age (y)	60.7 ± 10.3	60.7 ± 9.1
Age range	30.6 – 84.3	39.2 – 78.3
Gender (male, %)	57 (55.3%)	80 (66.1%)
Education (y)	16.6 ± 3.3	16.1 ± 3.0
UPDRS III OFF baseline	–	23.4 ± 10.1
UPDRS III OFF follow-up	–	25.8 ± 11.1
Duration T2 - T1 (y)	1.4 ± 0.5	1.4 ± 0.7

Table 2: **Abbreviations:** MCI = Mild Cognitive Impairment; UPDRS = Unified Parkinson’s Disease Rating Scale; PD = Parkinson’s disease. Values are expressed as mean \pm standard deviation. PD-non-MCI longitudinal sample is a subsample of the PD-non-MCI original sample that had longitudinal data and disease severity scores available.

4.3 Image acquisition and preprocessing

T1-weighted MRI images were obtained from PPMI that uses standardized acquisition parameters: repetition time = 2.3 s, echo time = 2.98 s, inversion time = 0.9 s, slice thickness = 1 mm, number of slices = 192, field of view = 256 mm, and matrix size = 256×256 . However, since PPMI is a multisite project there may be slight differences in the sites’ setup.

We processed images using FreeSurfer 7.3.1 instrumented with Fuzzy-libm to introduce controlled numerical perturbations. Each participant underwent 34 `recon-all` executions, extracting cortical thickness, surface area, and volumes. After quality control and exclusion of failed runs, we randomly selected 26 successful repetitions per subject to ensure balanced datasets for statistical analysis.

Longitudinal processing followed the standard FreeSurfer stream [RSRF12]: cross-sectional processing of both timepoints, followed by creation of an unbiased within-subject template [RF11] using robust registration [RRF10]. Downstream analyses used unperturbed FreeSurfer to prevent additional numerical perturbations.

4.4 Numerical Variability Assessment

We assessed FreeSurfer 7.3.1 numerical stability in cross-sectional and longitudinal contexts using the Numerical-Anatomical Variability Ratio (NAVR) and its relationship to statistical effect sizes.

4.4.1 Numerical-Anatomical Variability Ratio (NAVR)

To quantify computational stability relative to biological variation, we developed the Numerical-Anatomical Variability Ratio (NAVR). For each brain region, NAVR measures the ratio of measurement uncertainty arising from computational processes to natural inter-subject anatomical variation:

$$\text{NAVR} = \frac{\sigma_{\text{num}}}{\sigma_{\text{anat}}}$$

where σ_{num} represents numerical variability (measurement precision across MCA repetitions for individual subjects) and σ_{anat} represents anatomical variability (inter-subject differences within each repetition).

For each region of interest, measurements from n MCA repetitions across m subject-visit pairs form a data matrix $\mathcal{M}_{n \times m}$, where element $x_{i,j}$ represents the measurement for subject j in repetition i .

Numerical variability quantifies intra-subject measurement consistency:

$$\sigma_{\text{num}}^2 = \frac{1}{m} \sum_{j=1}^m \left[\frac{1}{n-1} \sum_{i=1}^n (x_{i,j} - \bar{x}_{\cdot,j})^2 \right] \quad (1)$$

Anatomical variability captures inter-subject differences:

$$\sigma_{\text{anat}}^2 = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{m-1} \sum_{j=1}^m (x_{i,j} - \bar{x}_{i,\cdot})^2 \right] \quad (2)$$

where $\bar{x}_{\cdot,j}$ and $\bar{x}_{i,\cdot}$ denote column and row means, respectively. Higher NAVR values indicate regions where computational uncertainty approaches or exceeds biological variation, potentially compromising the detection of true anatomical differences.

4.4.2 Relationship between NAVR and Effect Size Uncertainty

We derived the theoretical relationship between NAVR and Cohen's d variability to quantify how measurement uncertainty affects statistical effect sizes in group comparisons.

For a balanced two-group design with total sample size N , each observation decomposes as $X_{ij} = \mu_i + \varepsilon_{ij}^{(\text{anat})} + \varepsilon_{ij}^{(\text{num})}$, where μ_i represents the true group mean, $\varepsilon_{ij}^{(\text{anat})} \sim \mathcal{N}(0, \sigma_{\text{anat}}^2)$ captures anatomical variation, and $\varepsilon_{ij}^{(\text{num})} \sim \mathcal{N}(0, \sigma_{\text{num}}^2)$ represents numerical uncertainty.

The standard deviation of Cohen's d attributable to measurement error is:

$$\sigma_d = \frac{2}{\sqrt{N}} \cdot \text{NAVR} \quad (3)$$

This relationship emerges from error propagation analysis. The difference in group means has variance $\text{Var}(\bar{X}_1 - \bar{X}_2) = 4(\sigma_{\text{anat}}^2 + \sigma_{\text{num}}^2)/N$, with the numerical component contributing $4\sigma_{\text{num}}^2/N$. Since Cohen's d normalizes by the pooled standard deviation $\sqrt{\sigma_{\text{anat}}^2 + \sigma_{\text{num}}^2}$, the measurement error contribution becomes $\sigma_d = (2\sigma_{\text{num}}/\sqrt{N})/\sigma_{\text{anat}} = (2/\sqrt{N}) \cdot \text{NAVR}$.

This formula indicates that regions with $\text{NAVR} = 0.1$ contribute approximately $0.2/\sqrt{N}$ uncertainty to Cohen's d, while regions with $\text{NAVR} = 1.0$ contribute $2/\sqrt{N}$ uncertainty. The relationship provides a direct link between computational stability (NAVR) and statistical reliability in neuroimaging studies.

4.4.3 Cross-sectional Analysis

We extracted cortical and subcortical volumes, thickness, and surface areas from 26 MCA repetitions using the DK atlas. Numerical precision was quantified using significant digits, while measurement variability was assessed through standard deviations across repetitions.

Group comparisons between HC and PD used two-sample t-tests. Structural overlap was evaluated using the extended Sørensen-Dice coefficient:

$$\text{Dice}(A_1, A_2, \dots, A_n) = \frac{n |\bigcap_{i=1}^n A_i|}{\sum_{i=1}^n |A_i|}$$

4.4.4 Statistical Analysis

Longitudinal analyses examined correlations between brain metrics and UPDRS scores, plus group differences between PD and HC across two timepoints.

For subcortical regions, we assessed baseline volume-UPDRS correlations and group differences using partial correlation and ANCOVA. Longitudinal changes were analyzed using volume change rates: $(\text{volume}_2 - \text{volume}_1)/\text{volume}_1 \times 100$.

Vertex-wise cortical analyses examined thickness-UPDRS correlations and group differences at baseline and longitudinally. Thickness change rates were calculated as $(\text{thickness}_2 - \text{thickness}_1)/(\text{time}_2 - \text{time}_1)$ (mm/year). All analyses included age and sex covariates, with additional time-between-visits adjustment for longitudinal models. Cluster-wise permutation testing used $p < 0.05$ threshold, reporting cluster frequency across 26 MCA repetitions.

5 Data Availability

The data that support the findings of this study are available from the Parkinson’s Progression Markers Initiative (PPMI) database (www.ppmi-info.org/access-data-specimens/download-data), but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the PPMI.

6 Code Availability

The code used to conduct the analyses is available at [URL to be added upon publication].

7 Acknowledgements

The analyses were conducted on the Virtual Imaging Platform [[GLG⁺12](#)], which utilizes resources provided by the Biomed virtual organization within the European Grid Infrastructure (EGI). We extend our gratitude to Sorina Pop from CREATIS, Lyon, France, for her support.

References

- [BBD⁺21] Nikhil Bhagwat, Amadou Barry, Erin W Dickie, Shawn T Brown, Gabriel A Devenyi, Koji Hatano, Elizabeth DuPre, Alain Dagher, Mallar Chakravarty, Celia MT Greenwood, et al. Understanding the impact of preprocessing pipelines on neuroimaging cortical surface analyses. *GigaScience*, 10(1):giaa155, 2021.
- [BNHC⁺20] Rotem Botvinik-Nezer, Felix Holzmeister, Colin F Camerer, Anna Dreber, Juergen Huber, Magnus Johannesson, Michael Kirchler, Roni Iwanir, Jeanette A Mumford, R Alison Adcock, et al. Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, 582(7810):84–88, 2020.
- [DdOCP16] Christophe Denis, Pablo de Oliveira Castro, and Eric Petit. Verificarlo: checking floating point accuracy through monte carlo arithmetic. In *2016 IEEE 23nd Symposium on Computer Arithmetic (ARITH)*, 2016.
- [GHJ⁺12] Ed HBM Gronenschild, Petra Habets, Heidi IL Jacobs, Ron Mengelers, Nico Rozendaal, Jim Van Os, and Machteld Marcelis. The effects of freesurfer version, workstation type, and macintosh operating system version on anatomical volume and cortical thickness measurements. *PloS one*, 7(6):e38234, 2012.

- [GLG⁺12] Tristan Glatard, Carole Lartizien, Bernard Gibaud, Rafael Ferreira Da Silva, Germain Forestier, Frédéric Cervenansky, Martino Alessandrini, Hugues Benoit-Cattin, Olivier Bernard, Sorina Camarasu-Pop, et al. A virtual imaging platform for multi-modality medical image simulation. *IEEE transactions on medical imaging*, 32(1):110–118, 2012.
- [HPZ⁺23] Elizabeth Haddad, Fabrizio Pizzagalli, Alyssa H Zhu, Ravi R Bhatt, Tasfiya Islam, Iyad Ba Gari, Daniel Dixon, Sophia I Thomopoulos, Paul M Thompson, and Neda Jahanshad. Multisite test-retest reliability and compatibility of brain metrics derived from freesurfer versions 7.1, 6.0, and 5.3. *Human Brain Mapping*, 44(4):1515–1532, 2023.
- [Par97] Douglass Stott Parker. *Monte Carlo arithmetic: exploiting randomness in floating-point arithmetic*. Citeseer, 1997.
- [RF11] Martin Reuter and Bruce Fischl. Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage*, 57(1):19–21, 2011.
- [RRF10] Martin Reuter, H Diana Rosas, and Bruce Fischl. Highly accurate inverse consistent registration: a robust approach. *Neuroimage*, 53(4):1181–1196, 2010.
- [RSRF12] Martin Reuter, Nicholas J Schmansky, H Diana Rosas, and Bruce Fischl. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage*, 61(4):1402–1418, 2012.
- [SCF⁺21] Devan Sohier, Pablo De Oliveira Castro, François Févotte, Bruno Lathuilière, Eric Petit, and Olivier Jamond. Confidence intervals for stochastic arithmetic. *ACM Transactions on Mathematical Software (TOMS)*, 47(2):1–33, 2021.
- [SCKG21] Ali Salari, Yohan Chatelain, Gregory Kiar, and Tristan Glatard. Accurate simulation of operating system updates in neuroimaging using monte-carlo arithmetic. In *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Perinatal Imaging, Placental and Preterm Image Analysis: 3rd International Workshop, UNSURE 2021, and 6th International Workshop, PIPPI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings 3*, pages 14–23. Springer, 2021.

A Formula

A.1 Significant digits formula

We compute the number of significant bits \hat{s} with probability $p_s = 0.95$ and confidence $1 - \alpha_s = 0.95$ using the *Significant Digits* package¹ (version 0.2.0). *Significant Digits* implements the Centered Normality Hypothesis approach described in [SCF⁺21]:

$$\hat{s}_i = -\log_2 \left| \frac{\hat{\sigma}_i}{\hat{\mu}_i} \right| - \delta(n, \alpha_s, p_s),$$

where $\hat{\sigma}_i$ and $\hat{\mu}_i$ are the average and standard deviation over the repetitions, and

$$\delta(n, \alpha_s, p_s) = \log_2 \left(\sqrt{\frac{n-1}{\chi^2_{1-\alpha_s/2}}} \Phi^{-1} \left(\frac{p_s+1}{2} \right) \right) \quad (4)$$

is a penalty term for estimating \hat{s}_i with probability p_s and confidence level $1 - \alpha_s$ for a sample size n . Φ^{-1} is the inverse cumulative distribution of the standard normal distribution and χ^2 is the Chi-2 distribution with $n-1$ degrees of freedom.

B Cross-sectional Analysis

As a side result, the cross-sectional analysis measures the impact of numerical variability in FreeSurfer version 7.3.1 on the PPMI (Parkinson’s Progression Markers Initiative) cohort. This will involve comparing the estimation of structural MRI measures, including cortical and subcortical volumes, cortical thickness, and surface area. The goal is to assess the stability of these key metrics and determine how numerical variability may affect their reliability in clinical research.

B.1 Significant digits average across all subjects and regions

Table 3: Significant digits average across all subjects and regions.

Region	cortical thickness		surface area		cortical volume	
	lh	rh	lh	rh	lh	rh
G Ins lg and S cent ins	1.16 ± 0.19	1.14 ± 0.19	0.85 ± 0.12	0.72 ± 0.17	0.88 ± 0.13	0.84 ± 0.14
G and S cingul-Ant	1.58 ± 0.19	1.61 ± 0.18	1.13 ± 0.20	1.15 ± 0.21	1.11 ± 0.21	1.12 ± 0.19
G and S cingul-Mid-Ant	1.57 ± 0.17	1.59 ± 0.17	1.05 ± 0.20	1.08 ± 0.21	1.02 ± 0.19	1.05 ± 0.20
G and S cingul-Mid-Post	1.65 ± 0.17	1.62 ± 0.18	1.23 ± 0.20	1.18 ± 0.23	1.20 ± 0.21	1.14 ± 0.24
G and S frontomargin	1.42 ± 0.21	1.30 ± 0.20	1.07 ± 0.21	0.92 ± 0.16	1.00 ± 0.19	0.86 ± 0.16
G and S occipital inf	1.50 ± 0.18	1.50 ± 0.17	1.07 ± 0.16	1.05 ± 0.15	1.03 ± 0.16	1.04 ± 0.17
G and S paracentral	1.45 ± 0.23	1.45 ± 0.24	1.12 ± 0.15	1.16 ± 0.19	1.05 ± 0.17	1.09 ± 0.21
G and S subcentral	1.62 ± 0.18	1.62 ± 0.18	1.11 ± 0.14	1.13 ± 0.19	1.10 ± 0.15	1.14 ± 0.18
G and S transv frontopol	1.30 ± 0.25	1.34 ± 0.22	0.91 ± 0.16	0.91 ± 0.15	0.87 ± 0.19	0.92 ± 0.17
G cingul-Post-dorsal	1.51 ± 0.17	1.48 ± 0.18	0.98 ± 0.14	0.95 ± 0.17	0.99 ± 0.16	0.97 ± 0.19
G cingul-Post-ventral	1.20 ± 0.16	1.28 ± 0.18	0.77 ± 0.13	0.87 ± 0.15	0.75 ± 0.14	0.90 ± 0.16
G cuneus	1.41 ± 0.22	1.43 ± 0.22	1.24 ± 0.17	1.22 ± 0.19	1.19 ± 0.18	1.16 ± 0.19
G front inf-Opercular	1.66 ± 0.22	1.65 ± 0.21	1.09 ± 0.17	1.09 ± 0.20	1.11 ± 0.18	1.12 ± 0.21
G front inf-Orbital	1.41 ± 0.24	1.31 ± 0.24	0.78 ± 0.15	0.67 ± 0.16	0.81 ± 0.17	0.67 ± 0.17
G front inf-Triangul	1.56 ± 0.23	1.44 ± 0.24	1.03 ± 0.17	0.90 ± 0.22	1.06 ± 0.19	0.92 ± 0.23
G front middle	1.70 ± 0.23	1.62 ± 0.23	1.17 ± 0.23	0.97 ± 0.22	1.21 ± 0.24	0.99 ± 0.23
G front sup	1.78 ± 0.23	1.73 ± 0.23	1.37 ± 0.20	1.28 ± 0.22	1.39 ± 0.21	1.33 ± 0.22
G insular short	1.21 ± 0.23	1.06 ± 0.18	0.89 ± 0.13	0.65 ± 0.14	1.00 ± 0.16	0.89 ± 0.14
G oc-temp lat-fusifor	1.58 ± 0.19	1.58 ± 0.20	1.15 ± 0.17	1.13 ± 0.18	1.17 ± 0.18	1.15 ± 0.19

Continued on next page

¹<https://github.com/verificarlo/significantdigits>

Table 3: Significant digits average across all subjects and regions. (Continued)

Region	cortical thickness		surface area		cortical volume	
	lh	rh	lh	rh	lh	rh
G oc-temp med-Lingual	1.37 ± 0.23	1.35 ± 0.23	1.25 ± 0.18	1.26 ± 0.18	1.25 ± 0.19	1.26 ± 0.20
G oc-temp med-Parahip	1.41 ± 0.22	1.35 ± 0.21	1.02 ± 0.20	0.88 ± 0.22	1.02 ± 0.21	0.96 ± 0.19
G occipital middle	1.64 ± 0.18	1.66 ± 0.18	1.01 ± 0.17	1.10 ± 0.18	1.14 ± 0.18	1.15 ± 0.19
G occipital sup	1.55 ± 0.19	1.57 ± 0.20	1.17 ± 0.13	1.10 ± 0.15	1.15 ± 0.15	1.09 ± 0.16
G orbital	1.53 ± 0.24	1.43 ± 0.21	1.27 ± 0.18	1.08 ± 0.16	1.28 ± 0.18	1.22 ± 0.16
G pariet inf-Angular	1.69 ± 0.21	1.71 ± 0.18	1.11 ± 0.19	1.05 ± 0.21	1.10 ± 0.20	1.06 ± 0.21
G pariet inf-Supramar	1.71 ± 0.21	1.74 ± 0.19	1.15 ± 0.20	1.19 ± 0.26	1.16 ± 0.21	1.20 ± 0.27
G parietal sup	1.73 ± 0.22	1.68 ± 0.22	1.16 ± 0.22	1.06 ± 0.24	1.21 ± 0.23	1.09 ± 0.25
G postcentral	1.62 ± 0.26	1.57 ± 0.30	1.21 ± 0.21	1.23 ± 0.27	1.23 ± 0.21	1.26 ± 0.26
G precentral	1.56 ± 0.30	1.51 ± 0.35	1.26 ± 0.18	1.25 ± 0.23	1.29 ± 0.22	1.26 ± 0.28
G precuneus	1.70 ± 0.19	1.70 ± 0.21	1.22 ± 0.19	1.19 ± 0.21	1.24 ± 0.19	1.20 ± 0.23
G rectus	1.35 ± 0.21	1.29 ± 0.22	0.96 ± 0.15	0.96 ± 0.15	0.95 ± 0.15	0.96 ± 0.16
G subcallosal	1.07 ± 0.13	1.02 ± 0.13	0.55 ± 0.15	0.61 ± 0.11	0.57 ± 0.15	0.67 ± 0.13
G temp sup-G T transv	1.42 ± 0.23	1.41 ± 0.21	0.97 ± 0.16	0.88 ± 0.12	1.00 ± 0.18	0.93 ± 0.13
G temp sup-Lateral	1.63 ± 0.23	1.66 ± 0.22	1.02 ± 0.19	1.26 ± 0.16	1.22 ± 0.20	1.28 ± 0.18
G temp sup-Plan polar	1.31 ± 0.23	1.22 ± 0.21	0.81 ± 0.17	0.69 ± 0.17	0.81 ± 0.19	0.79 ± 0.16
G temp sup-Plan tempo	1.61 ± 0.20	1.62 ± 0.19	1.07 ± 0.18	1.06 ± 0.15	1.03 ± 0.18	1.07 ± 0.14
G temporal inf	1.54 ± 0.20	1.56 ± 0.21	1.08 ± 0.20	1.15 ± 0.18	1.09 ± 0.22	1.19 ± 0.19
G temporal middle	1.64 ± 0.19	1.72 ± 0.18	1.15 ± 0.19	1.27 ± 0.17	1.21 ± 0.20	1.32 ± 0.18
Lat Fis-ant-Horizont	1.33 ± 0.20	1.39 ± 0.20	0.74 ± 0.20	0.84 ± 0.19	0.65 ± 0.19	0.76 ± 0.19
Lat Fis-ant-Vertical	1.32 ± 0.21	1.22 ± 0.23	0.66 ± 0.21	0.58 ± 0.25	0.58 ± 0.21	0.49 ± 0.23
Lat Fis-post	1.62 ± 0.18	1.67 ± 0.15	1.09 ± 0.14	1.09 ± 0.14	1.00 ± 0.14	1.02 ± 0.14
Pole occipital	1.42 ± 0.21	1.46 ± 0.22	1.14 ± 0.14	1.23 ± 0.16	1.06 ± 0.18	1.14 ± 0.17
Pole temporal	1.34 ± 0.27	1.42 ± 0.24	1.06 ± 0.21	1.11 ± 0.20	1.00 ± 0.24	1.08 ± 0.23
S calcarine	1.45 ± 0.21	1.42 ± 0.21	1.25 ± 0.22	1.21 ± 0.21	1.21 ± 0.18	1.17 ± 0.17
S central	1.60 ± 0.25	1.61 ± 0.24	1.37 ± 0.25	1.33 ± 0.30	1.32 ± 0.21	1.28 ± 0.26
S cingul-Marginalis	1.66 ± 0.20	1.65 ± 0.21	1.18 ± 0.18	1.19 ± 0.22	1.07 ± 0.18	1.06 ± 0.22
S circular insula ant	1.38 ± 0.22	1.37 ± 0.22	0.98 ± 0.16	0.96 ± 0.18	0.86 ± 0.15	0.84 ± 0.16
S circular insula inf	1.45 ± 0.17	1.50 ± 0.17	1.09 ± 0.18	1.09 ± 0.17	0.98 ± 0.15	0.98 ± 0.15
S circular insula sup	1.58 ± 0.18	1.54 ± 0.17	1.12 ± 0.13	1.02 ± 0.15	1.01 ± 0.13	0.94 ± 0.14
S collat transv ant	1.33 ± 0.20	1.40 ± 0.19	0.93 ± 0.19	0.96 ± 0.21	0.83 ± 0.19	0.85 ± 0.19
S collat transv post	1.31 ± 0.19	1.30 ± 0.18	0.81 ± 0.17	0.73 ± 0.16	0.72 ± 0.18	0.65 ± 0.17
S front inf	1.61 ± 0.19	1.58 ± 0.20	1.13 ± 0.24	1.06 ± 0.25	1.06 ± 0.23	0.99 ± 0.25
S front middle	1.46 ± 0.19	1.41 ± 0.20	0.90 ± 0.23	0.85 ± 0.22	0.81 ± 0.24	0.70 ± 0.23
S front sup	1.68 ± 0.19	1.63 ± 0.19	1.09 ± 0.22	0.96 ± 0.25	1.01 ± 0.21	0.87 ± 0.24
S interm prim-Jensen	1.22 ± 0.28	1.33 ± 0.21	0.61 ± 0.30	0.65 ± 0.23	0.53 ± 0.29	0.57 ± 0.22
S intrapariet and P trans	1.72 ± 0.18	1.71 ± 0.19	1.14 ± 0.22	1.10 ± 0.25	1.11 ± 0.21	1.03 ± 0.23
S oc-temp lat	1.41 ± 0.19	1.47 ± 0.18	0.83 ± 0.17	0.96 ± 0.19	0.73 ± 0.17	0.86 ± 0.19
S oc-temp med and Lingual	1.58 ± 0.17	1.61 ± 0.17	1.15 ± 0.19	1.16 ± 0.19	1.08 ± 0.18	1.10 ± 0.17
S oc middle and Lunatus	1.49 ± 0.19	1.47 ± 0.19	1.00 ± 0.20	0.96 ± 0.21	0.91 ± 0.19	0.87 ± 0.20
S oc sup and transversal	1.60 ± 0.17	1.59 ± 0.18	1.17 ± 0.18	1.12 ± 0.20	1.10 ± 0.18	1.06 ± 0.18
S occipital ant	1.40 ± 0.19	1.47 ± 0.17	0.67 ± 0.15	0.85 ± 0.15	0.60 ± 0.16	0.76 ± 0.15
S orbital-H Shaped	1.50 ± 0.18	1.48 ± 0.19	1.16 ± 0.26	1.14 ± 0.28	1.14 ± 0.23	1.14 ± 0.22
S orbital lateral	1.23 ± 0.20	1.21 ± 0.19	0.69 ± 0.19	0.65 ± 0.20	0.59 ± 0.19	0.56 ± 0.18
S orbital med-olfact	1.32 ± 0.17	1.20 ± 0.17	0.92 ± 0.17	0.60 ± 0.13	0.85 ± 0.14	0.63 ± 0.13
S parieto occipital	1.70 ± 0.19	1.69 ± 0.21	1.32 ± 0.19	1.33 ± 0.19	1.24 ± 0.18	1.27 ± 0.19
S pericallosal	1.21 ± 0.13	1.24 ± 0.14	0.98 ± 0.25	1.04 ± 0.26	0.86 ± 0.21	0.92 ± 0.21
S postcentral	1.74 ± 0.20	1.73 ± 0.22	1.20 ± 0.27	1.20 ± 0.31	1.13 ± 0.25	1.13 ± 0.27
S precentral-inf-part	1.64 ± 0.22	1.62 ± 0.28	1.21 ± 0.24	1.19 ± 0.26	1.13 ± 0.24	1.10 ± 0.29
S precentral-sup-part	1.57 ± 0.24	1.51 ± 0.31	1.09 ± 0.22	1.09 ± 0.23	1.03 ± 0.21	1.00 ± 0.24
S suborbital	1.27 ± 0.16	1.05 ± 0.20	0.74 ± 0.16	0.52 ± 0.22	0.61 ± 0.15	0.43 ± 0.20
S subparietal	1.53 ± 0.17	1.50 ± 0.17	1.04 ± 0.21	0.98 ± 0.23	0.95 ± 0.20	0.88 ± 0.21

Continued on next page

Table 3: Significant digits average across all subjects and regions. (Continued)

Region	cortical thickness		surface area		cortical volume	
	lh	rh	lh	rh	lh	rh
S temporal inf	1.45 ± 0.16	1.46 ± 0.15	0.86 ± 0.17	0.94 ± 0.17	0.76 ± 0.16	0.83 ± 0.16
S temporal sup	1.73 ± 0.17	1.75 ± 0.15	1.29 ± 0.22	1.33 ± 0.19	1.23 ± 0.20	1.24 ± 0.18
S temporal transverse	1.32 ± 0.19	1.30 ± 0.18	0.82 ± 0.19	0.77 ± 0.19	0.73 ± 0.18	0.68 ± 0.18

Table 4: Standard-deviation average across all subjects and regions.

Region	cortical thickness (mm)		surface area (mm ²)		cortical volume (mm ³)	
	lh	rh	lh	rh	lh	rh
G Ins lg and S cent ins	0.09 ± 0.04	0.10 ± 0.05	26.80 ± 12.53	34.88 ± 19.32	81.42 ± 48.62	88.35 ± 46.33
G and S cingul-Ant	0.03 ± 0.02	0.03 ± 0.02	51.54 ± 39.32	63.44 ± 42.94	145.54 ± 108.64	173.81 ± 99.91
G and S cingul-Mid-Ant	0.03 ± 0.02	0.03 ± 0.02	34.57 ± 24.10	35.63 ± 22.48	89.14 ± 52.30	95.86 ± 57.77
G and S cingul-Mid-Post	0.02 ± 0.02	0.02 ± 0.02	24.28 ± 19.46	30.54 ± 27.37	65.07 ± 53.03	85.59 ± 75.73
G and S frontomargin	0.04 ± 0.03	0.05 ± 0.03	31.83 ± 21.28	31.61 ± 13.38	94.56 ± 64.13	97.65 ± 46.48
G and S occipital inf	0.03 ± 0.02	0.03 ± 0.02	39.02 ± 19.81	32.57 ± 15.92	119.36 ± 56.10	99.07 ± 46.18
G and S paracentral	0.03 ± 0.02	0.04 ± 0.02	33.65 ± 17.96	29.67 ± 26.32	109.16 ± 66.28	94.71 ± 72.49
G and S subcentral	0.03 ± 0.02	0.03 ± 0.02	33.87 ± 19.69	31.97 ± 21.67	104.20 ± 61.20	89.71 ± 60.11
G and S transv frontopol	0.06 ± 0.05	0.05 ± 0.03	28.12 ± 12.80	40.99 ± 16.84	104.12 ± 65.82	134.29 ± 65.41
G cingul-Post-dorsal	0.04 ± 0.03	0.04 ± 0.04	19.23 ± 14.88	17.77 ± 12.57	64.32 ± 44.09	59.57 ± 35.92
G cingul-Post-ventral	0.07 ± 0.05	0.06 ± 0.05	15.66 ± 17.15	12.99 ± 22.56	50.17 ± 27.85	38.86 ± 36.83
G cuneus	0.03 ± 0.02	0.03 ± 0.02	35.26 ± 30.33	40.96 ± 42.52	84.71 ± 74.31	101.9 ± 97.02
G front inf-Opercular	0.03 ± 0.02	0.03 ± 0.02	32.65 ± 25.00	35.08 ± 34.37	105.47 ± 78.45	108.38 ± 98.3
G front inf-Orbital	0.05 ± 0.04	0.06 ± 0.04	21.20 ± 11.01	25.25 ± 10.37	71.05 ± 46.15	86.33 ± 38.23
G front inf-Triangul	0.03 ± 0.03	0.04 ± 0.03	30.29 ± 19.17	37.08 ± 32.23	91.18 ± 65.35	116.92 ± 104.8
G front middle	0.02 ± 0.02	0.03 ± 0.02	92.58 ± 75.11	126.22 ± 77.76	273.22 ± 234.35	403.59 ± 245.74
G front sup	0.02 ± 0.02	0.02 ± 0.02	97.83 ± 73.49	113.70 ± 97.83	299.95 ± 206.52	339.13 ± 278.46
G insular short	0.09 ± 0.05	0.11 ± 0.05	29.04 ± 15.99	47.90 ± 23.20	98.87 ± 58.14	112.37 ± 56.78
G oc-temp lat-fusifor	0.03 ± 0.02	0.03 ± 0.03	37.79 ± 21.25	41.78 ± 26.17	128.12 ± 71.17	146.71 ± 94.52
G oc-temp med-Lingual	0.04 ± 0.02	0.04 ± 0.02	51.64 ± 39.25	48.19 ± 36.79	122.30 ± 98.61	118.59 ± 98.07
G oc-temp med-Parahip	0.05 ± 0.03	0.06 ± 0.03	37.22 ± 36.52	49.76 ± 69.91	143.85 ± 100.75	158.91 ± 121.6
G occipital middle	0.02 ± 0.02	0.02 ± 0.02	47.90 ± 22.84	52.98 ± 30.88	138.60 ± 79.51	148.24 ± 95.59
G occipital sup	0.03 ± 0.02	0.03 ± 0.02	26.48 ± 14.00	34.90 ± 21.23	68.90 ± 41.73	91.73 ± 53.70
G orbital	0.03 ± 0.02	0.04 ± 0.02	39.63 ± 27.95	63.63 ± 27.59	126.67 ± 98.09	162.43 ± 78.39
G pariet inf-Angular	0.02 ± 0.02	0.02 ± 0.01	57.33 ± 29.55	77.98 ± 47.01	179.81 ± 96.06	236.38 ± 142.78
G pariet inf-Supramar	0.02 ± 0.02	0.02 ± 0.01	62.02 ± 46.17	54.59 ± 60.03	192.54 ± 148.64	170.62 ± 190.83
G parietal sup	0.02 ± 0.01	0.02 ± 0.01	62.42 ± 52.63	65.13 ± 50.00	161.05 ± 166.21	177.73 ± 159.43
G postcentral	0.03 ± 0.03	0.03 ± 0.03	42.75 ± 36.99	39.75 ± 47.18	106.60 ± 82.38	96.85 ± 114.98
G precentral	0.04 ± 0.04	0.04 ± 0.05	42.85 ± 36.80	47.68 ± 45.38	138.71 ± 137.38	164.16 ± 191.76
G precuneus	0.02 ± 0.02	0.02 ± 0.02	49.44 ± 43.95	49.64 ± 40.76	145.37 ± 137.81	149.86 ± 128.30
G rectus	0.05 ± 0.03	0.06 ± 0.04	32.20 ± 15.60	27.50 ± 12.79	106.76 ± 51.68	91.10 ± 49.31
G subcallosal	0.08 ± 0.02	0.09 ± 0.03	51.68 ± 17.71	34.17 ± 12.84	110.85 ± 37.44	73.96 ± 31.21
G temp sup-G T transv	0.04 ± 0.04	0.04 ± 0.03	15.35 ± 8.69	13.73 ± 5.19	42.31 ± 28.07	37.23 ± 14.97
G temp sup-Lateral	0.03 ± 0.02	0.03 ± 0.02	38.43 ± 27.47	29.34 ± 15.20	147.99 ± 97.89	111.92 ± 61.68
G temp sup-Plan polar	0.07 ± 0.05	0.08 ± 0.04	32.33 ± 18.14	47.67 ± 23.04	119.94 ± 70.10	124.65 ± 58.27
G temp sup-Plan tempo	0.03 ± 0.02	0.03 ± 0.02	24.87 ± 16.10	20.93 ± 9.79	71.54 ± 39.15	55.07 ± 22.84
G temporal inf	0.03 ± 0.02	0.03 ± 0.02	63.81 ± 45.99	52.37 ± 28.88	230.73 ± 163.99	180.27 ± 100.37
G temporal middle	0.03 ± 0.02	0.02 ± 0.01	54.58 ± 37.80	46.74 ± 27.32	184.08 ± 135.35	155.87 ± 83.42
Lat Fis-ant-Horizont	0.04 ± 0.03	0.04 ± 0.02	15.78 ± 07.02	16.00 ± 9.18	37.09 ± 18.17	37.24 ± 22.79
Lat Fis-ant-Vertical	0.05 ± 0.03	0.06 ± 0.06	19.09 ± 8.70	16.93 ± 7.9	47.83 ± 23.08	42.18 ± 19.39
Lat Fis-post	0.02 ± 0.02	0.02 ± 0.02	24.88 ± 13.15	30.93 ± 13.87	56.61 ± 30.98	71.03 ± 31.70

Continued on next page

Table 4: Standard-deviation average across all subjects and regions. (Continued)

Region	cortical thickness (mm)		surface area (mm ²)		cortical volume (mm ³)	
	lh	rh	lh	rh	lh	rh
Pole occipital	0.03 ± 0.02	0.03 ± 0.02	40.70 ± 18.96	56.64 ± 36.50	113.61 ± 66.05	157.76 ± 95.97
Pole temporal	0.06 ± 0.04	0.05 ± 0.03	46.23 ± 30.07	43.02 ± 26.78	251.05 ± 187.99	215.93 ± 147.79
S calcarine	0.03 ± 0.02	0.03 ± 0.02	44.93 ± 39.69	47.41 ± 45.06	77.13 ± 67.45	83.23 ± 65.81
S central	0.02 ± 0.02	0.02 ± 0.02	45.58 ± 48.37	52.54 ± 69.94	80.26 ± 74.65	92.02 ± 109.44
S cingul-Marginalis	0.02 ± 0.02	0.02 ± 0.02	21.30 ± 17.52	25.74 ± 22.14	53.01 ± 41.86	71.65 ± 58.25
S circular insula ant	0.05 ± 0.04	0.05 ± 0.04	15.81 ± 10.25	18.35 ± 12.25	46.26 ± 34.95	53.18 ± 33.38
S circular insula inf	0.04 ± 0.03	0.03 ± 0.02	33.02 ± 19.40	27.57 ± 16.37	92.16 ± 47.87	73.13 ± 35.95
S circular insula sup	0.03 ± 0.03	0.03 ± 0.02	35.16 ± 16.49	36.41 ± 20.50	91.74 ± 44.20	89.54 ± 47.94
S collat transv ant	0.05 ± 0.04	0.04 ± 0.03	31.84 ± 18.98	31.78 ± 17.89	95.76 ± 52.02	97.81 ± 51.0
S collat transv post	0.04 ± 0.03	0.04 ± 0.02	22.36 ± 9.36	24.44 ± 10.34	49.00 ± 22.33	52.38 ± 23.3
S front inf	0.02 ± 0.02	0.02 ± 0.02	50.05 ± 48.10	55.04 ± 53.15	122.15 ± 115.86	137.13 ± 139.79
S front middle	0.03 ± 0.02	0.03 ± 0.02	51.70 ± 33.07	91.25 ± 60.17	137.58 ± 98.08	282.75 ± 181.72
S front sup	0.02 ± 0.01	0.02 ± 0.02	70.65 ± 53.57	88.18 ± 73.29	189.44 ± 133.17	242.99 ± 183.7
S interm prim-Jensen	0.07 ± 0.09	0.05 ± 0.04	22.91 ± 15.67	26.85 ± 14.81	55.28 ± 35.58	62.20 ± 32.78
S intrapariet and P trans	0.02 ± 0.02	0.02 ± 0.01	65.99 ± 57.12	83.26 ± 77.54	133.09 ± 111.32	186.66 ± 155.46
S oc-temp lat	0.04 ± 0.03	0.03 ± 0.02	37.74 ± 16.79	33.31 ± 16.25	101.88 ± 46.87	90.78 ± 43.89
S oc-temp med and Lingual	0.02 ± 0.02	0.02 ± 0.02	43.42 ± 29.10	40.87 ± 26.30	105.48 ± 57.26	93.32 ± 50.19
S oc middle and Lunatus	0.03 ± 0.02	0.03 ± 0.02	28.69 ± 15.64	29.71 ± 19.00	61.89 ± 33.18	65.79 ± 40.53
S oc sup and transversal	0.02 ± 0.02	0.02 ± 0.02	24.28 ± 14.75	30.11 ± 20.51	54.26 ± 32.40	65.35 ± 38.76
S occipital ant	0.04 ± 0.02	0.03 ± 0.02	40.53 ± 15.92	29.35 ± 13.03	98.81 ± 42.71	76.07 ± 35.41
S orbital-H Shaped	0.03 ± 0.02	0.03 ± 0.02	30.77 ± 27.50	30.63 ± 24.64	75.75 ± 60.88	69.25 ± 45.60
S orbital lateral	0.05 ± 0.03	0.05 ± 0.03	20.80 ± 10.56	28.47 ± 13.86	50.60 ± 26.90	66.88 ± 32.94
S orbital med-olfact	0.05 ± 0.03	0.06 ± 0.03	22.67 ± 9.39	52.03 ± 15.39	51.77 ± 21.69	96.61 ± 38.19
S parieto occipital	0.02 ± 0.02	0.02 ± 0.03	29.51 ± 23.78	32.84 ± 34.50	68.41 ± 46.21	72.54 ± 68.14
S pericallosal	0.04 ± 0.02	0.04 ± 0.02	44.29 ± 80.37	44.16 ± 60.70	78.45 ± 149.00	74.66 ± 86.88
S postcentral	0.02 ± 0.02	0.02 ± 0.02	59.33 ± 63.12	57.23 ± 76.08	132.12 ± 121.05	118.64 ± 135.52
S precentral-inf-part	0.02 ± 0.03	0.03 ± 0.04	32.56 ± 41.72	35.05 ± 38.33	82.93 ± 113.93	95.11 ± 120.79
S precentral-sup-part	0.03 ± 0.04	0.04 ± 0.05	32.51 ± 24.59	33.81 ± 26.76	74.56 ± 59.33	81.98 ± 69.50
S suborbital	0.05 ± 0.03	0.09 ± 0.08	37.26 ± 17.06	27.92 ± 11.25	101.77 ± 51.24	71.72 ± 29.02
S subparietal	0.03 ± 0.03	0.03 ± 0.03	29.36 ± 27.35	38.95 ± 35.41	70.32 ± 55.27	101.75 ± 74.08
S temporal inf	0.03 ± 0.02	0.03 ± 0.02	60.38 ± 30.96	41.64 ± 17.77	156.80 ± 75.30	106.77 ± 42.99
S temporal sup	0.02 ± 0.02	0.02 ± 0.02	86.27 ± 60.78	83.72 ± 54.81	217.81 ± 141.29	220.72 ± 129.73
S temporal transverse	0.04 ± 0.03	0.05 ± 0.03	17.51 ± 9.28	14.56 ± 6.46	40.39 ± 21.69	36.01 ± 17.13

Table 5: Significant digits average across all subjects and regions.

Region	Subcortical volume	Region	Subcortical volume
3rd-Ventricle	1.44 ± 0.20	CC Anterior	1.34 ± 0.36
4th-Ventricle	1.32 ± 0.20	CC Central	1.08 ± 0.38
5th-Ventricle	13.51 ± 4.98	CC Mid Anterior	1.20 ± 0.39
Brain-Stem	1.70 ± 0.21	CC Mid Posterior	1.13 ± 0.36
CSF	1.14 ± 0.21	CC Posterior	1.38 ± 0.47
Left-Accumbens-area	0.87 ± 0.17	Right-Accumbens-area	0.98 ± 0.16
Left-Amygdala	1.12 ± 0.16	Right-Amygdala	1.22 ± 0.18
Left-Caudate	1.56 ± 0.21	Right-Caudate	1.50 ± 0.26
Left-Cerebellum-Cortex	1.88 ± 0.23	Right-Cerebellum-Cortex	1.87 ± 0.25
Left-Cerebellum-White-Matter	1.24 ± 0.24	Right-Cerebellum-White-Matter	1.25 ± 0.28
Left-Hippocampus	1.47 ± 0.18	Right-Hippocampus	1.53 ± 0.19
Optic-Chiasm	0.76 ± 0.23		

Continued on next page

Table 5: Significant digits average across all subjects and regions. (Continued)

Region	Subcortical volume	Region	Subcortical volume
Left-Inf-Lat-Vent	0.82 ± 0.23	Right-Inf-Lat-Vent	0.88 ± 0.27
Left-Lateral-Ventricle	1.88 ± 0.25	Right-Lateral-Ventricle	1.83 ± 0.28
Left-Pallidum	1.24 ± 0.20	Right-Pallidum	1.21 ± 0.21
Left-Putamen	1.47 ± 0.23	Right-Putamen	1.49 ± 0.28
Left-Thalamus	1.41 ± 0.22	Right-Thalamus	1.42 ± 0.22
Left-VentralDC	1.40 ± 0.16	Right-VentralDC	1.39 ± 0.15
Left-WM-hypointensities	15.22 ± 0.00	Right-WM-hypointensities	15.22 ± 0.00
Left-choroid-plexus	0.81 ± 0.16	Right-choroid-plexus	0.83 ± 0.16
Left-non-WM-hypointensities	15.22 ± 0.00	Right-non-WM-hypointensities	15.22 ± 0.00
Left-vessel	0.46 ± 0.63	Right-vessel	0.23 ± 0.89
Optic-Chiasm	0.76 ± 0.23		

Table 6: Standard-deviation average across all subjects and regions.

Region	Subcortical volume (mm ³)	Region	Subcortical volume (mm ³)
3rd-Ventricle	25.18 ± 34.65	CC Anterior	23.67 ± 29.16
4th-Ventricle	38.47 ± 23.49	CC Central	23.89 ± 23.90
5th-Ventricle	0.09 ± 0.43	CC Mid Anterior	19.86 ± 24.04
Brain-Stem	187.47 ± 103.98	CC Mid Posterior	21.69 ± 29.83
CSF	38.22 ± 49.09	CC Posterior	31.59 ± 53.10
Left-Accumbens-area	24.50 ± 9.50	Right-Accumbens-area	20.89 ± 8.86
Left-Amygdala	49.73 ± 20.47	Right-Amygdala	43.96 ± 24.00
Left-Caudate	40.85 ± 28.91	Right-Caudate	51.88 ± 45.72
Left-Cerebellum-Cortex	313.86 ± 227.78	Right-Cerebellum-Cortex	338.10 ± 271.17
Left-Cerebellum-White-Matter	380.83 ± 224.68	Right-Cerebellum-White-Matter	379.10 ± 329.94
Left-Hippocampus	57.14 ± 40.36	Right-Hippocampus	51.70 ± 39.33
Left-Inf-Lat-Vent	31.66 ± 16.10	Right-Inf-Lat-Vent	28.75 ± 15.28
Left-Lateral-Ventricle	88.18 ± 185.84	Right-Lateral-Ventricle	121.41 ± 578.96
Left-Pallidum	49.21 ± 30.03	Right-Pallidum	52.09 ± 37.88
Left-Putamen	69.42 ± 51.72	Right-Putamen	73.75 ± 96.90
Left-Thalamus	124.64 ± 81.18	Right-Thalamus	124.52 ± 111.74
Left-VentralDC	65.09 ± 30.33	Right-VentralDC	65.07 ± 27.96
Left-WM-hypointensities	15.22 ± 0.00	Right-WM-hypointensities	0.00 ± 0.00
Left-choroid-plexus	43.74 ± 22.99	Right-choroid-plexus	44.79 ± 25.13
Left-non-WM-hypointensities	0.00 ± 0.00	Right-non-WM-hypointensities	0.00 ± 0.00
Left-vessel	6.42 ± 4.91	Right-vessel	7.92 ± 5.38
Optic-Chiasm	13.11 ± 6.56		

C Longitudinal Analysis

C.1 Vertex-wise Analysis

C.1.1 Z-Test Results

C.1.2 Group analysis

The following table summarizes significant regions identified in the group analysis using a Z-test ($\alpha = 0.05$). Each region was statistically significant in only one MCA repetition, highlighting a lack of replicability.

Table 7: Significant regions identified using a Z-test ($\alpha = 0.05$) for the group analysis.

Region	Size (mm ²)	MNI X	MNI Y	MNI Z	Max	Frequency
Baseline						
R postcentral	844.98	23.8	-36.4	57.2	-2.964	1 / 25
Longitudinal						
L inferior parietal	3337.65	-56.6	-43.6	-18.1	2.8572	1 / 25
R parstriangularis	3265.87	44.4	35.8	3.7	2.8129	1 / 25

C.1.3 Correlation analysis

The following table presents significant regions identified in the correlation analysis using a Z-test ($\alpha = 0.05$). Notably, only two regions were found to be significant in two different MCA repetitions, with variability in their estimated size and MNI coordinates.

Table 8: Significant regions identified using a Z-test ($\alpha = 0.05$) for the correlation analysis.

Region	Size (mm ²)	MNI X	MNI Y	MNI Z	Max	Frequency
Baseline						
L postcentral	1006.20 ± 39.27	-37.0	-33.8 / -34	62.5 / 61.7	-4.7429 / -3.7576	2/25
L precentral	1011.32	-18.2	-19.8	70.1	-4.2896	1/25
R lateral occipital	997.20 ± 43.08	14.2 / 20.7	-99.8 / -96.5	1.9 / -13.1	-2.9967 / -3.1263	2/25
R pericalcarine	2085.19	10.8	-74.9	4.4	-3.000	1/25

Table 9: Summary of executions failure and excluded subjects. To standardize the sample, we keep 25 repetitions per subject/visits pair. Subject/visit pairs with less than 25 repetitions were excluded which is 12 subjects.

Stage	Number of rejected repetitions	Total number of repetitions
Cluster failure	1246 (5.80%)	21488
FreeSurfer failure	68 (0.33%)	21488
QC failure	319 (1.48%)	21488
Total	1633 (7.60%)	21488

D NAVR

D.1 NAVR maps

D.2 Consistency results

D.2.1 Consistency of statistical tests

D.2.2 Distribution of statistical tests coefficients

Status	Cohort	HC	PD-non-MCI	PD-MCI
Before QC	n	106	181	29
	Age (y)	60.6 ± 10.2	61.7 ± 9.6	67.7 ± 7.7
	Age range	30.6 – 84.3	36.3 – 83.3	49.9 – 80.5
	Gender (male, %)	58 (54.7%)	119 (65.7%)	–
After QC	Education (y)	16.6 ± 3.3	15.9 ± 2.9	–
	n	103	175	27
	Age (y)	60.7 ± 10.3	61.4 ± 9.5	67.8 ± 7.9
	Age range	30.6 – 84.3	36.3 – 79.9	49.9 – 80.5
After MCI exclusion	Gender (male, %)	57 (55.3%)	114 (65.1%)	20 (74.1%)
	Education (y)	16.6 ± 3.3	15.9 ± 2.9	15.0 ± 3.5
	n	103	121	–
	Age (y)	60.7 ± 10.3	60.7 ± 9.1	–
After MCI exclusion	Age range	30.6 – 84.3	39.2 – 78.3	–
	Gender (male, %)	57 (55.3%)	80 (66.1%)	–
	Education (y)	16.6 ± 3.3	16.1 ± 3.0	–
	UPDRS III OFF baseline	–	23.4 ± 10.1	–
After MCI exclusion	UPDRS III OFF follow-up	–	25.8 ± 11.1	–
	Duration T2 - T1 (y)	1.4 ± 0.5	1.4 ± 0.7	–

Abbreviations: MCI = Mild Cognitive Impairment; UPDRS = Unified Parkinson's Disease Rating Scale; PD = Parkinson's disease. Descriptive statistics before and after quality control (QC). Values are expressed as mean ± standard deviation. PD-non-MCI longitudinal sample is a subsample of the PD-non-MCI original sample that had longitudinal data and disease severity scores available.

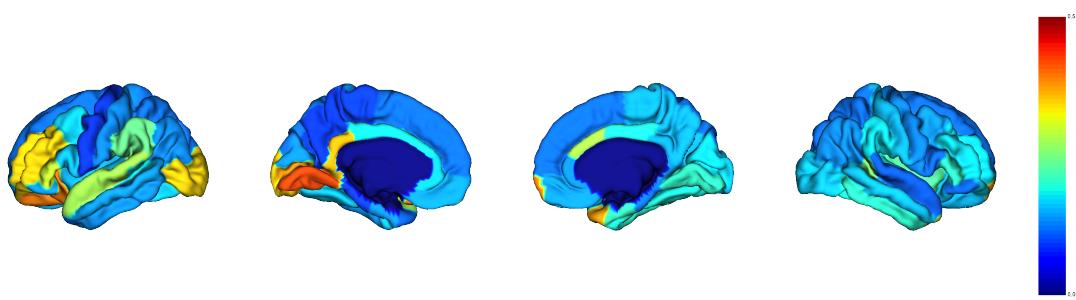


Figure 10: NAVR maps for cortical surface area. The maps show the average NAVR values across all subjects for each cortical region. The color scale indicates the NAVR value, with warmer colors indicating higher NAVR values.

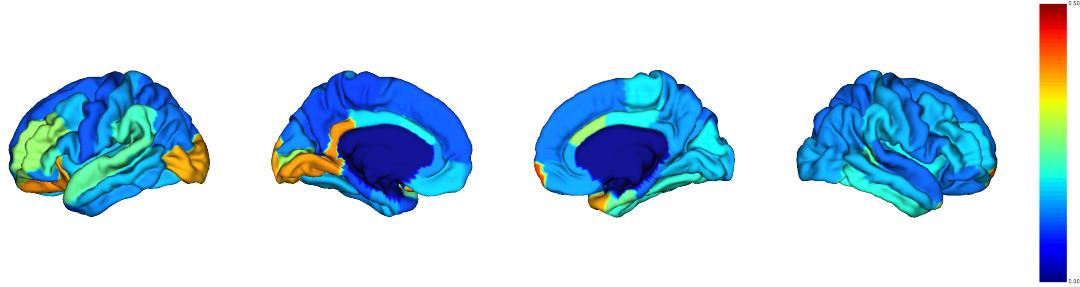


Figure 11: NAVR maps for cortical volume. The maps show the average NAVR values across all subjects for each cortical region. The color scale indicates the NAVR value, with warmer colors indicating higher NAVR values.



Figure 12: Consistency of statistical tests for cortical area across all subjects and regions. The plot shows the percentage of subjects for which the statistical test was significant ($\alpha = 0.05$) for each region. The consistency varies across regions, with some regions showing higher consistency than others.

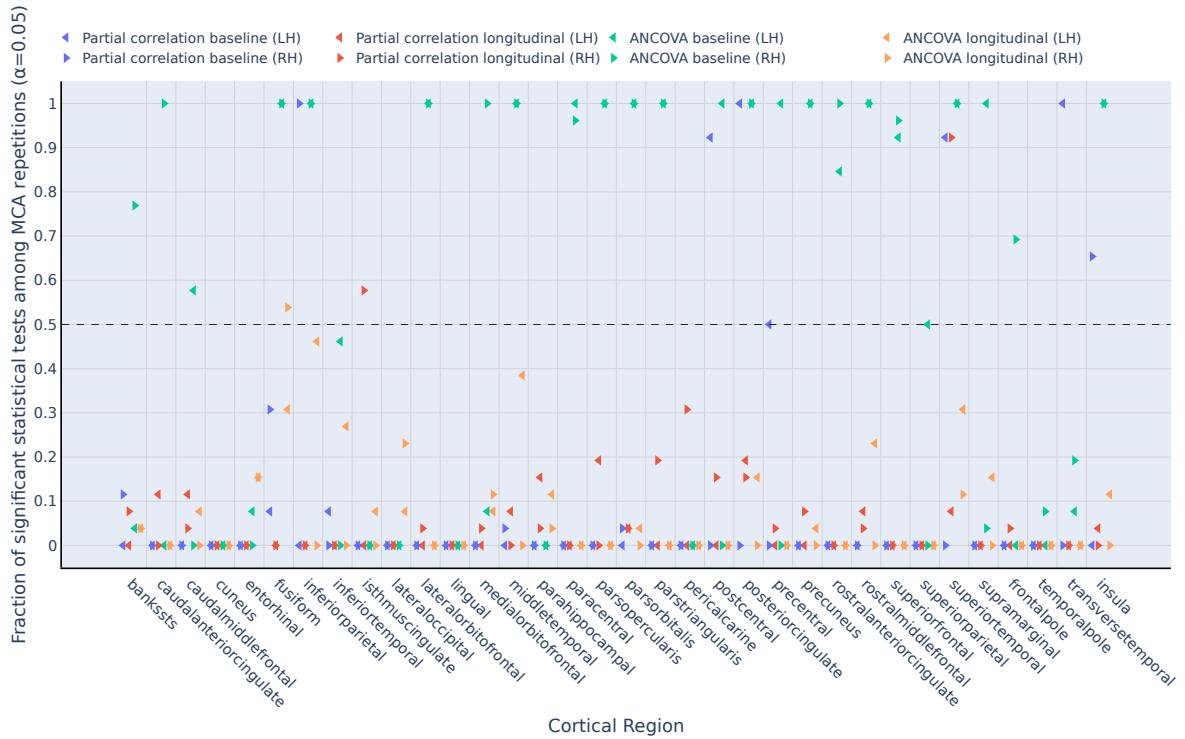
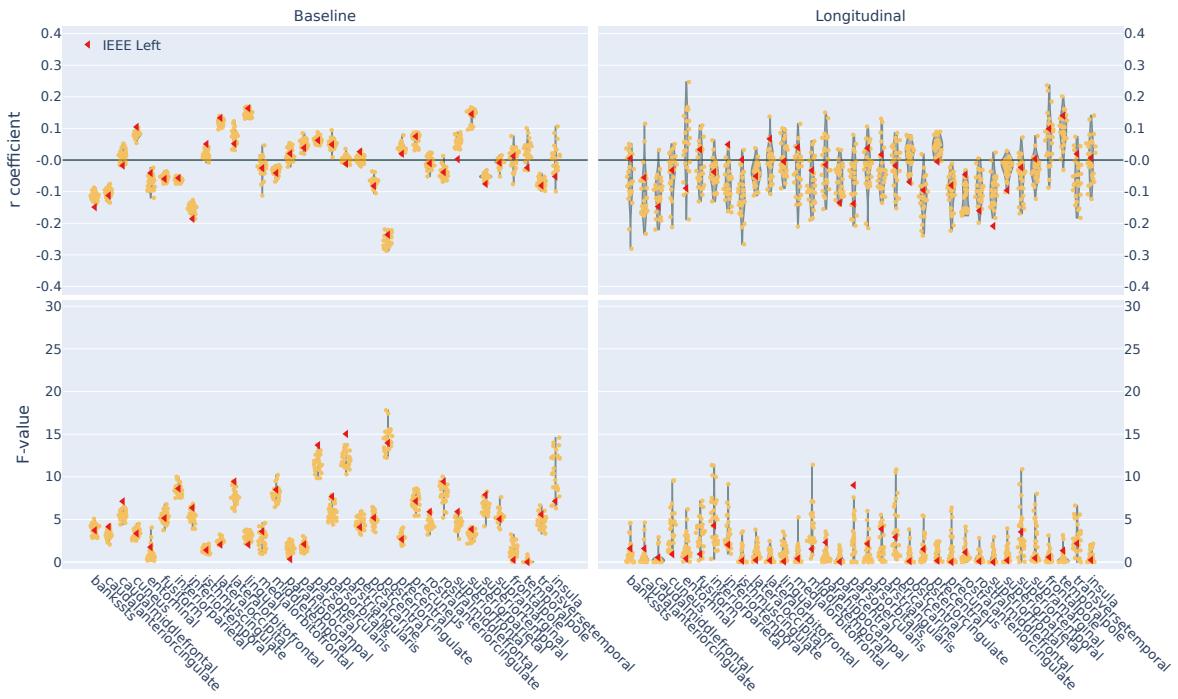
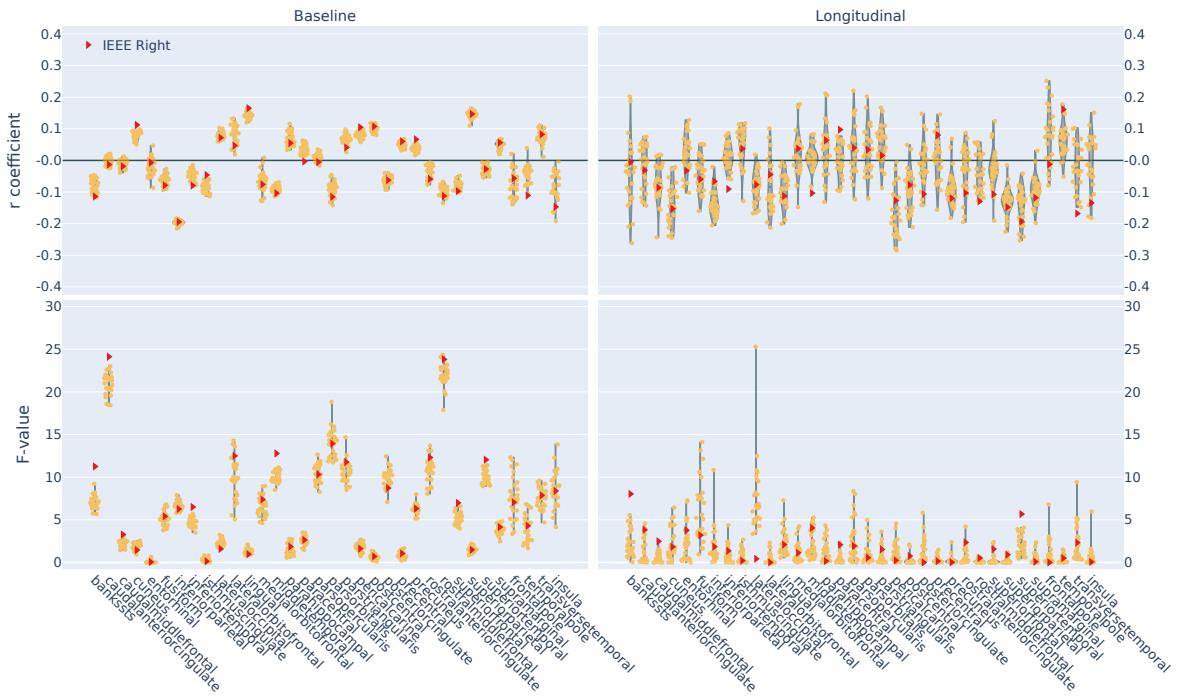


Figure 13: Consistency of statistical tests for cortical volume across all subjects and regions. The plot shows the percentage of subjects for which the statistical test was significant ($\alpha = 0.05$) for each region. The consistency varies across regions, with some regions showing higher consistency than others.

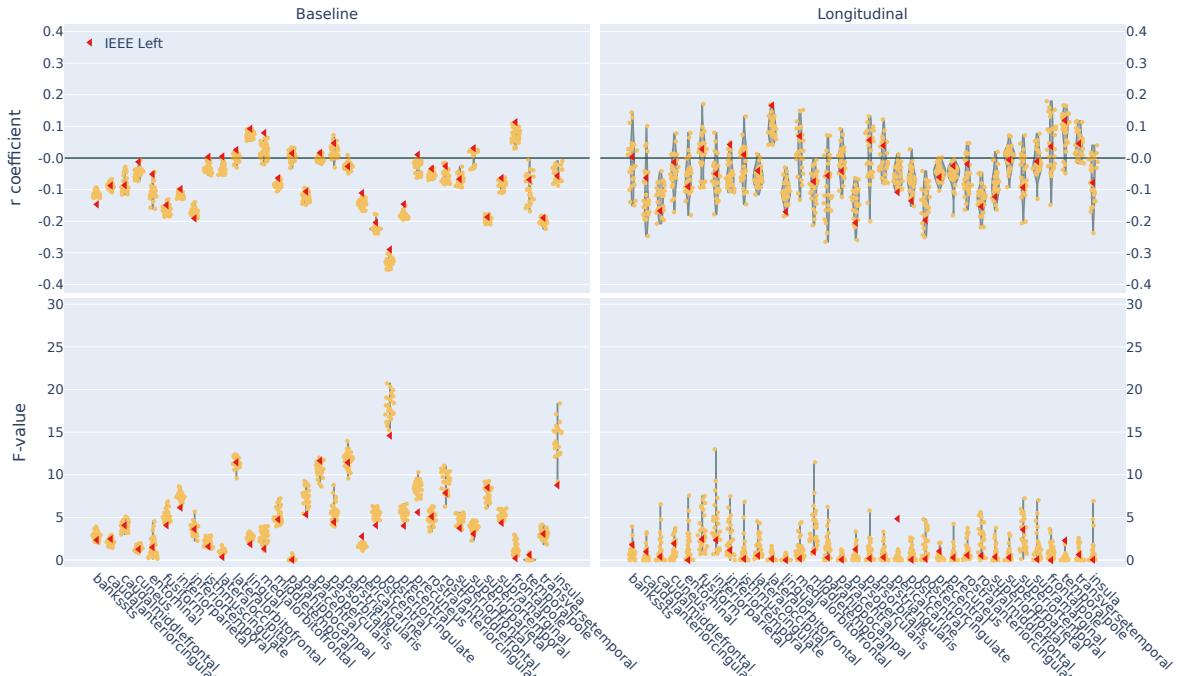


(a) Left hemisphere

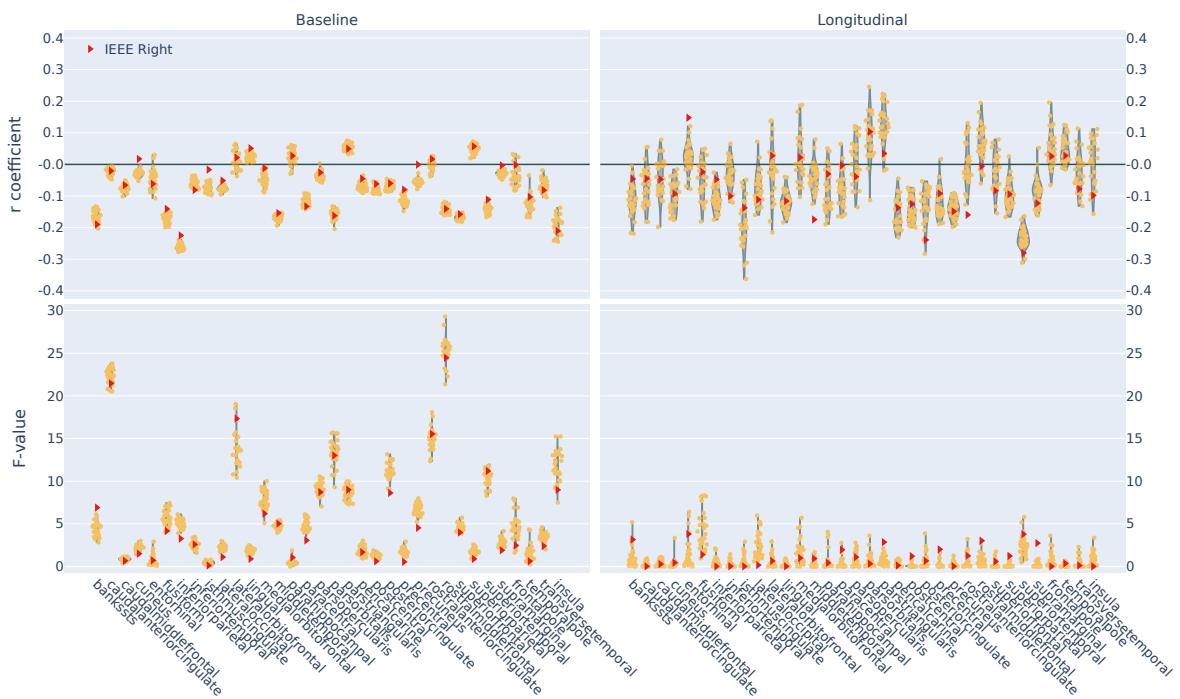


(b) Right hemisphere

Figure 14: Distribution of partial correlation coefficients for cortical area across all subjects and regions. Red triangles indicate the IEEE-754 run for reference. The distribution shows the variability in the coefficients, with some regions exhibiting higher consistency than others.



(a) Left hemisphere



(b) Right hemisphere

Figure 15: Distribution of partial correlation coefficients for cortical volume across all subjects and regions. Red triangles indicate the IEEE-754 run for reference. The distribution shows the variability in the coefficients, with some regions exhibiting higher consistency than others.