nature research

Corresponding author(s):	Ross Markello & Bratislav Misic
Last updated by author(s):	Nov 3, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our <u>Editorial Policies</u> and the <u>Editorial Policy Checklist</u>.

~				
51	ta	ŤΙ	ロスキ	ICS

For	all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.
n/a	Confirmed
	\square The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
	🔀 A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
	The statistical test(s) used AND whether they are one- or two-sided Only common tests should be described solely by name; describe more complex techniques in the Methods section.
	A description of all covariates tested
	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
	For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
\times	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
\boxtimes	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
	Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated
	Our web collection on <u>statistics for biologists</u> contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Data used in this study were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (https://www.ppmi-info.org), accessed in March of 2018.

Data analysis

All code used for data processing, analysis, and figure generation is available on GitHub (https://github.com/netneurolab/markello_ppmisnf). Analyses were conducted in Python 3.6.7; the exact computing environment used to perform all analyses is packaged into a Singularity container, available on OSF (https://osf.io/h6jwx/).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a <u>data availability statement</u>. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data used in this study were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (https://www.ppmi-info.org), accessed in March of 2018. Processed data used in the reported analyses are available on GitHub via https://github.com/netneurolab/markello_ppmisnf or via Zenodo at https://doi.org/10.5281/zenodo.3731252.

Field-spe	ecific re	porting		
Please select the or	ne below that is	the best fit for your research. If you are not sure, read the appropriate sections before making your selection.		
Life sciences	В	ehavioural & social sciences Ecological, evolutionary & environmental sciences		
For a reference copy of t	he document with a	ill sections, see <u>nature.com/documents/nr-reporting-summary-flat.pdf</u>		
Life scier	nces stu	ıdy design		
All studies must dis	close on these	points even when the disclosure is negative.		
Sample size		e calculations were performed for the current analysis as we relied on the one-of-a-kind, open-access PPMI database. The 273 lealthy controls and 186 patients with Parkinson's disease) were those who met inclusion criteria (see below).		
Data exclusions	enrollment. Sub	s://www.ppmi-info.org/study-design/research-documents-and-sops/ for documentation regarding exclusion criteria for subject subjects were excluded from data analysis if they (1) did not have at least one T1-weighted MRI scan or (2) were missing >20% of any data modality.		
Replication	the PPMI for a s	not replicated as the authors are unaware of other Parkinson's disease datasets that contain all modalities of data collected by similarly-sized (or larger) cohort (i.e., structural MRI, dopamine active transporter scans, biospecimen assays from fluid, and broad clinical-behavioral assessments).		
Randomization		allocated into experimental groups based on PD diagnosis made during study enrollment. Refer to https://www.ppmi-info.org/esearch-documents-and-sops/ for documentation describing diagnostic procedures.		
Blinding	Investigators we	ors were not blind to experimental group as the analyses were focused solely on one of the groups (i.e., PD patients).		
We require information system or method list	on from authors a ed is relevant to	Decific materials, systems and methods bout some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.		
Materials & exp				
n/a Involved in th	· ·	n/a Involved in the study		
Antibodies Antibodies Eukaryotic		ChIP-seq Flow cytometry		
	ogy and archaeol			
Animals and other organisms				
Clinical dat	а			
Dual use re	esearch of concer	ı		
Human rese	arch parti	cipants		
Policy information	about <u>studies ir</u>	volving human research participants		
Population chara	cteristics	Healthy individuals with no history of neurological illness (n = 87; ages 30-83, 32F/55M) and individuals diagnosed with Parkinson's disease (n = 186; ages 36-82, 78F/108M).		
Recruitment	Refer to https://www.ppmi-info.org/study-design/research-documents-and-sops/ for documentation describing recruitment of subjects for the PPMI.			
Ethics oversight	thics oversight Central ethics oversight for the PPMI was provided by UCSF and additional oversight by individual research sites participate in the PPMI.			

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type

No task was run during scanning as all MRI scans used were structural images.

Design specifications	Scan time varied based on imaging sequence (see Acquisition, below).		
Behavioral performance measure	res No behavioral measures were collected during scanning.		
Acquisition			
Imaging type(s)	Structural (T1-weighted, T2-weighted, proton-density, FLAIR)		
Field strength	1.5T & 3T		
Sequence & imaging parameters	Imaging sequence varied by site in the PPMI; for a full list of sites and sequences please refer to https://www.ppmi-info.org/.		
Area of acquisition	Whole-brain scans were used.		
Diffusion MRI Used	Not used ■ Not used		
Preprocessing			
Preprocessing software	he Advanced Normalization Tools (ANTs) software v2.2.0 was used for all image processing.		
Normalization	T1w images were combined with other available neuroimages (T2-weighted, proton-density, and FLAIR images) across all subject timepoints to create a temporally unbiased, subject-specific template. A standard template was non-linearly registered to the subject template and used to remove non-brain tissue. The remaining subject template brain tissue was segmented into six classes (cerebrospinal fluid, cortical gray matter, white matter, subcortical gray matter, brain stem, and cerebellum) using ANTs joint label fusion with a group of fifteen expertly annotated and labeled atlases from the OASIS dataset. Finally, the segmented subject template was used to performing the same nonlinear registration, brain extraction, and segmentation procedures on the T1w images from each timepoint.		
Normalization template	We used the MNI-ICBM152 non-linear 2009c asymmetric template for normalization.		
Noise and artifact removal	T1-weighted structural images were corrected for signal intensity non-uniformity with N4 bias correction.		
Volume censoring	No volume censoring was performed; all data were anatomical.		
Statistical modeling & infere	ce		
Model type and settings	Cortical thickness and subcortical nuclei volumes were estimated for each subject from their processed T1w brain images. Where applicable, subjects were modeled as random effects.		
Effect(s) tested	One-way ANOVAs were used to examine differences in imaging-based features of interest between derived PD patient biotypes.		
Specify type of analysis: Wi	ole brain 🔲 ROI-based 🔀 Both		
Anatomical location(s) ROI-based analyses were conducted using the Lausanne anatomical atlas, which is a subdivision of the Desikan-Killiany atlas (Cammoun et al., 2012).			
Statistic type for inference (See <u>Eklund et al. 2016</u>)	No voxel-wise or cluster corrections were performed as whole-brain images were used solely for computation of a single, summary-statistic PD atrophy score.		
Correction	FDR corrections (Benjamini-Hochberg) were used where appropriate to correct for multiple comparisons.		
Models & analysis			
n/a Involved in the study Functional and/or effective connectivity			
Graph analysis			
Multivariate modeling or predictive analysis			
Multivariate modeling and predictive analysis No multivariate modeling was performed on the MRI data itself; however, the data were used data fusion analyses (i.e., similarity network fusion).			