

## 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 [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

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

- |                                     |                                     |                                                                                                                                                                                                                                                            |
|-------------------------------------|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( <i>n</i> ) for each experimental group/condition, given as a discrete number and unit of measurement                                                                                                                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>                                                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of all covariates tested                                                                                                                                                                                                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons                                                                                                                                        |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | 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) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | 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<br><i>Give P values as exact values whenever suitable.</i>                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings                                                                                                                                                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes                                                                                                                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | 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 [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection Data was collected from groups that had already used such data in their studies

Data analysis For data analysis in the study Freesurfer 6.0.0, IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp, R and R studio R Core Team, 2017 and in particular (but not only) the following packages: MASS, braingraph, igraph, factominer, factoextra, cocor.

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 [data availability statement](#). 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

The datasets as extracted upon Freesurfer segmentation and parcellation will be shared upon publication and uploaded on the Open Science Framework page of the project. Data will be shared pre-harmonisation and post-harmonisation as well.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	There was no prior power calculation. However, we wanted to reach a number of patients meaningfully higher than that reported in studies on this topic.
Data exclusions	Participants' scan that did not present good quality (e.g. motion) were discarded before the pre-processing, together with participants who did not meet our criteria for cognition, or healthy controls.
Replication	We conducted a sensitivity analysis with a subsample (N=146), repeating the same group analyses as those conducted for the main sample and additional correlational analyses with more specific scores of hallucinations.
Randomization	Participants were allocated in the groups in each original study. We retained the original classifications.
Blinding	Being a study on pre-collected data, blinding was not necessary.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input type="checkbox"/> Human research participants
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data
<input type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involved in the study
<input type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used *Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.*

Validation *Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.*

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s) *State the source of each cell line used.*

Authentication *Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*

Mycoplasma contamination *Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*

Commonly misidentified lines (See [ICLAC](#) register) *Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

## Palaeontology and Archaeology

Specimen provenance	Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).
Specimen deposition	Indicate where the specimens have been deposited to permit free access by other researchers.
Dating methods	If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.
<input type="checkbox"/> Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.	
Ethics oversight	Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

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

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.
Wild animals	Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.
Field-collected samples	For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.
Ethics oversight	Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

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

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	493 participants (193 F), of which 135 were PD-VH. Each individual study had matched their participants for age, gender, disease onset, Mini mental state examination MMSE, UPDRS-III and levodopa equivalent daily dose (LED), except for a study where MMSE score was lower in PD-VH, a study with UPDRS-III scores higher in PD-VH, and two studies where gender was not matched. We also included the unpublished data in separate ANOVAs to check that groups were matched, in meta-analyses and in an ANOVA including the whole mega-analysis sample. While the ANOVAs and the meta-analysis demonstrated we have good matching on the criteria, the mega-analysis ANOVA shows that there is a difference of 2.19 years in age [F (1,491) = 6.56, p = .01] (PD-VH = 67.85, SD = 7.74; 62 F, and PD-noVH = 65.66, SD = 8.71; 131 F) and there is a greater proportion of females in the PD-VH group ( $\chi^2 = 3.585$ , p = .06).
Recruitment	Participants were recruited at the time of the study at each site.
Ethics oversight	Ethical approval for the study (LRS-19/20-17680) was given on the 25/03/2020 by King's College London Research Ethics Office, Psychiatry, Nursing and Midwifery (PNM) Research Ethics Panel.

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

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.
Study protocol	Note where the full trial protocol can be accessed OR if not available, explain why.
Data collection	Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.
Outcomes	Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.

## Dual use research of concern

Policy information about [dual use research of concern](#)

### Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Public health
<input checked="" type="checkbox"/>	<input type="checkbox"/> National security
<input checked="" type="checkbox"/>	<input type="checkbox"/> Crops and/or livestock
<input checked="" type="checkbox"/>	<input type="checkbox"/> Ecosystems
<input checked="" type="checkbox"/>	<input type="checkbox"/> Any other significant area

### Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Demonstrate how to render a vaccine ineffective
<input checked="" type="checkbox"/>	<input type="checkbox"/> Confer resistance to therapeutically useful antibiotics or antiviral agents
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enhance the virulence of a pathogen or render a nonpathogen virulent
<input checked="" type="checkbox"/>	<input type="checkbox"/> Increase transmissibility of a pathogen
<input checked="" type="checkbox"/>	<input type="checkbox"/> Alter the host range of a pathogen
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enable evasion of diagnostic/detection modalities
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enable the weaponization of a biological agent or toxin
<input checked="" type="checkbox"/>	<input type="checkbox"/> Any other potentially harmful combination of experiments and agents

## ChIP-seq

### Data deposition

- ☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

#### Data access links

May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

#### Files in database submission

Provide a list of all files available in the database submission.

#### Genome browser session

(e.g. [UCSC](#))

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

### Methodology

#### Replicates

Describe the experimental replicates, specifying number, type and replicate agreement.

#### Sequencing depth

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

#### Antibodies

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

#### Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

#### Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

#### Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

## Flow Cytometry

### Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation

*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*

Instrument

*Identify the instrument used for data collection, specifying make and model number.*

Software

*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*

Cell population abundance

*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*

Gating strategy

*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

- ☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Magnetic resonance imaging

### Experimental design

Design type

Structural MRI study on pre-collected data investigating differences in cortical thickness and surface area in patients with Parkinson's disease and with and without visual hallucinations.

Design specifications

n/a

Behavioral performance measures

At the original site, different measures of hallucinations have been collected (Neuropsychiatric Inventory, or UPDRS, or NEVHI), Parkinson's severity information (UPDRS-III), cognition (MMSE), disease onset, levodopa equivalent daily dose.

### Acquisition

Imaging type(s)

structural

Field strength

mixed: each study had their own acquisition

Sequence & imaging parameters

mixed: each study had their own acquisition

Area of acquisition

whole brain

Diffusion MRI

☐ Used

☒ Not used

### Preprocessing

Preprocessing software

Freesurfer 6.0.0. with standard pre-processing pipeline recon-all. Freesurfer pre-processing was carried out on a high performance computing infrastructure using Slurm.

Normalization

Normalised to the Talairach template.

Normalization template

We used Freesurfer's fsaverage template.

Noise and artifact removal

Upon pre-processing, the mean and standard deviation of thickness was used to identify participants for which the process could have not been successful and careful manual inspection and, if needed, troubleshooting, were carried out.

Volume censoring

*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

### Statistical modeling & inference

Model type and settings

To explore group differences, we used a GLM (multivariate analysis of covariance), then we carried out Principal Component

Analyses, Regression models with receptor density maps from independent sources, a sensitivity analysis and structural covariance analyses.

Effect(s) tested

Differences between hallucinators and non hallucinators.

Specify type of analysis: ☐ Whole brain ☒ ROI-based ☐ Both

Anatomical location(s)

We used Destrieux and Desikan parcellations. This approach was necessary as multi-site harmonisation with the methodology used (empirical Bayes with Combat) was possible only on parcellated data.

Statistic type for inference  
(See [Eklund et al. 2016](#))

vertex-wise

Correction

FWE correction for multiple comparisons was used for the group analysis, for the graph theoretical analysis we used FDR and permutations.

## Models & analysis

n/a

Involved in the study

☐ ☐ Functional and/or effective connectivity

☐ ☒ Graph analysis

☐ ☒ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis

The dependent variable was presence or absence of visual hallucinations, and we tested this on structural covariance networks, we used binarised graphs, group-level. As node summaries we used efficiency, betweenness centrality and degree to determine hubs.

Multivariate modeling and predictive analysis

*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*