nature research

Corresponding author(s): Richard E. Daws

Last updated by author(s): Feb 3, 2022

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	
☐	☒	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. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	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 *d*, Pearson's *r*), 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	No software was used in data collection.
Data analysis	fMRI data preprocessing was conducted using FSL, AFNI, Freesurfer abd ANTs, All analyses and data visualizations were conducted in Matlab R2020a. Code from the Brain Connectivity Toolbox (BCT) (brain-connectivity-toolbox.net) was used for the fMRI network analysis. All analysis and data visualization codes will be made available at https://github.com/rdaws/psilodep.

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 data presented in each figure are available at https://github.com/rdaws/psilodep. The raw MRI data are available upon reasonable request to RCH.

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

Life sciences study design

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

Sample size	Study1: This feasibility open-label trial recruited 19 patients with treatment resistant unipolar depression. The sample size is due to the primary aims of feasibility and safety associated with this trial. Study2: This phase 2, double-blind, RCT recruited 59 patients with major unipolar depression. The sample size is was chosen to allow between trial-arm comparisons and was sufficient for detecting medium to large effect sizes.
Data exclusions	Study1: Three patients were excluded for excessive fMRI head-motion. This resulted in a final analytical sample of n=16. Study2: Of the 29 randomly assigned to the escitalopram-arm: 1 patient did not complete the follow-up due to Covid 19 UK 'lockdown', 4 patients discontinued the intervention due to adverse side effects and 3 patients were excluded due to excessive fMRI head-motion. Of the 30 randomly assigned to the psilocybin-arm, 2 patients did not complete follow-up due to Covid 19 Uk 'lockdown', 1 patient discontinued taking daily placebo tablets and 5 patients were excluded due to excessive fMRI head motion. Together, this resulted in a final analytical sample of n=21 in the escitalopram-arm and n=22 in the psilocybin-arm.
Replication	In both studies, psilocybin therapy was associated with significant decreases in brain network modularity. This robust replication was further bolstered by significant patient correlation analyses. In both studies, the reductions in modularity that followed psilocybin therapy correlated with long-term improvements in depression severity. Across studies, the secondary analyses of network-level cartography did not replicate; however, the inferences from each study converged on the same brain networks whose organisation changes in a consistent manner across studies. Moreover, the changes that followed psilocybin therapy were predicted by, and consistent with, previously reported functional abnormalities in depression. While consistent, these effects were observed with a 'static' approach in study 1 and with a 'dynamic' approach in study 2. The difference between trials may reflect significant differences in depression symptom severity, trial designs and/or fMRI acquisition parameters.
Randomization	Study1: All patients underwent the same protocol. No randomization was required for this open-label feasibility trial. All analyses depended on within-subject changes that followed the intervention. Supplemental analyses indicate that the findings are robust to confounding factors, including fMRI head-motion. Study2: A random number generator was used to randomly assign patients to the psilocybin-arm or escitalopram-arm.
Blinding	Study1: In this open-label trial, the dosage and compound were known to both the patients and investigators. Study2: In this double-blind RCT, the patients and investigators were blind to the trial-arm, psilocybin therapy dosage (a negligible 1mg vs 25mg) and daily capsule contents (escitalopram vs placebo). More specifically, patients in the psilocybin-arm received 2 x 25mg psilocybin therapy sessions and daily placebo tablets. Patients in the escitalopram-arm received 2 x 1mg psilocybin therapy sessions and daily escitalopram tablets.

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
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Study1: Mean age=42.75, SD=10.15, 4/16 female. Study2: Psilocybin-arm - Mean age 40.9, SD=10.1, 6/21 female. Escitalopram-arm - Mean age=44.5, SD =11.0, 8/22 female. For both trials, eligibility required a general practitioner (GP) confirmed diagnosis of unipolar major depressive disorder (MDD) (16+ on the 21-item Hamilton Depression Rating scale [HAM-D]). The open-label trial had the additional criteria of treatment-resistant depression, as no improvement despite multiple courses of antidepressant medication (mean = 4.6 ± 2.6 past medications, range = 2-11).
Recruitment	For both studies, exclusion criteria were: Immediate family or personal history of psychosis, risky physical health condition (physician-assessed), history of serious suicide attempts, positive pregnancy test and MRI contraindications. Study 2 had the additional exclusion criteria of selective serotonin reuptake inhibitor (SSRI) contraindications or previous escitalopram use. Eligible patients undertook telephone screening interviews, provided written informed consent and their mental and physical medical histories were thoroughly evaluated. It is possible that expectancy effects could have influenced the present results. For example, an approximate minority of 1/4 to 1/3 of patients reporting having prior experience of psychedelics.
Ethics oversight	Both studies were conducted at the NIHR Imperial Clinical Research Facility and received Imperial College London Sponsorship, NHS research and Imperial college Joint Research and Compliance Office (JRCO) ethical approval, Health Research Authority and MHRA approval. This work was done under a UK Home Office Schedule 1 Drug License.

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	Study1: UKRI MRC - MR/J00460X/1 Study2: Clinicaltrials.gov - NCT03429075
Study protocol	Study1: Carhart-Harris et al. (2016), doi:10.1016/S2215-0366(16)30065-7, gtr.ukri.org/projects?ref=MR%2FJ00460X%2F1 Study2: Carhart-Harris et al. (2021), doi:10.1056/NEJMoa2032994, clinicaltrials.gov/ct2/show/NCT03429075?term=NCT03429075&draw=2&rank=1
Data collection	For both studies, visits occurred at the NIHR Imperial Clinical Research Facility. Study1 recruitment and data collection: April-2015 to April-2016. Study2 recruitment and data collection: January-2019 to March-2020.
Outcomes	For both studies, Beck depression inventory (BDI) scores and resting-state fMRI global brain network modularity were the primary measures of interest. BDI was well-suited for testing the primary hypothesis due to its emphasis on the cognitive features of depression and it is a reliable index of broader subjective symptomatology - mappable to brain function. Brain network modularity is an elegant composite measure of the brain's functional community structure. Abnormally modular, or segregated, network organization has previously been observed in depression patients and this measure correlates with symptom severity. In contrast, the acute action of psilocybin promotes a global decrease in network modularity (high integration) in healthy adults. Moreover, psilocybin therapy has been seen to deliver rapid, substantial and sustained reductions in depression severity. Therefore, it was hypothesized that the efficacy of psilocybin therapy for depression is underpinned by a sub-acute 'carryover' effect, and predicted that this would be observed as a long-term reduction in depression severity that follows, and correlates with, a decrease in network modularity. Changes in fMRI network modularity scores between the baseline and follow-up scans were assessed with at the group-level with paired t-tests. Pearson correlation analyses were conducted between the changes in modularity and the BDI severity scores. The secondary analyses focused on network-level cartography. These analyses provided a deeper, finer-grained, understanding of the changes in brain function that follow psilocybin therapy. Network metrics were assessed at the group-level with paired t-tests and Pearson correlation analyses were conducted between the changes in network flexibility and the BDI severity scores.

Magnetic resonance imaging

Experimental design

Design type	Resting-state blood oxygen level dependent fMRI (eyes closed)
Design specifications	Study 1: Two rs-fMRI scans per subject, each with 240 measurements (TR=2000ms). Study 2: Two rs-fMRI scans per subject, each with 480 measurements (TR=1250ms).

Behavioral performance measures	None.

Acquisition

Imaging type(s)	Both studies: Functional MRI.
Field strength	Both studies: 3T Siemens Tim Trio
Sequence & imaging parameters	fMRI sequences. Study 1: 12 channel head coil, TR=2000ms, TE=31ms, 36 axial slices, flip angle=80 degrees, bandwidth=2298Hz/pixel, GRAPPA acceleration=2). Study 2: 32-channel head coil, TR = 1250 ms, TE = 30 ms, 44 axial slices, flip angle = 70 degrees, bandwidth = 2232 Hz/pixel, GRAPPA acceleration=2).
Area of acquisition	Both studies: Whole brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	The FMRIB Software Library (FSL), Analysis of Functional NeuroImages (AFNI), Freesurfer and Advanced Normalisation Tools (ANTs) packages were used to preprocess the data from both studies. Specifically, the following pre-processing stages were performed: 1) Removal of the first three volumes; 2) de-spiking (3dDespike, AFNI); 3) slice time correction (3dTshift, AFNI); 4) motion correction (3dvolreg, AFNI) by registering each volume to the volume most similar, in the least-squares sense, to all others; 5) brain extraction (BET, FSL). After normalization and scrubbing, 6mm full width at half maximum (FWHM) gaussian spatial smoothing (3dBlurInMask, AFNI); 0.01 to 0.08Hz band-pass filtering (3dFourier, AFNI) and linear and quadratic de-trending (3dDetrend, AFNI) was applied.
Normalization	fMRI volumes were co-registered to the anatomical scans using a rigid body registration (BBR, FSL). Non-linear registration to the 2mm MNI brain was then applied to the co-registered volumes (Symmetric Normalization (SyN), ANTS);
Normalization template	MNI305 2mm brain
Noise and artifact removal	Voxelwise nuisance regression was conducted with the 6 realignment motion regressors and 3 tissue signal regressors (ventricles (Freesurfer, eroded in 2mm space), draining veins (DV) (FSL's CSF minus Freesurfer's Ventricles, eroded in 1mm space) and local white matter (WM) (FSL's WM minus Freesurfer's subcortical grey matter (GM) structures, eroded in 2mm space). Regarding local WM regression, AFNI's 3dLocalstat was used to calculate the mean local WM time-series for each voxel, using a 25mm radius sphere centred on each voxel.
Volume censoring	Scrubbing: Using a framewise displacement (FD) threshold of 0.5 mm, scrubbed volumes were replaced with the mean of the neighbouring volumes.

Statistical modeling & inference

Model type and settings	Univariate composite scores of brain connectivity were compared between sessions and correlated with clinical outcomes. Analyses were corrected for multiple comparisons using the false discovery rate (FDR) correction.
Effect(s) tested	rs-fMRI functional connectivity composite scores (e.g., network modularity) were compared between sessions and correlated with clinical scores.

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

Anatomical location(s)	A publicly available functional atlas of the human cerebral cortex was used to segment the entire cortex into 100 regions of interest. These ROIs were used in the functional connectivity analyses. https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal
Statistic type for inference (See Eklund et al. 2016)	Statistical comparisons were made with composite scores of the functional connectivity matrices (e.g. network modularity). No task-activation analyses were conducted.
Correction	False discovery rate (FDR) correction was applied to correlations between functional network composite scores and clinical outcomes measures.

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	Pearson correlation was used to estimate functional connectivity. This was conducted between pair of ROIs to construct functional connectivity matrices for each subject and scanning session.
Graph analysis	Network modularity, and related measures, were estimated using a community detection algorithm applied to the positive functional connectivity weights. Measures were controlled for near-degeneracy and also normalized with a random reshuffling permutation approach. Modularity measures were estimated for each subject and scan before being compared at the group-level and in the individual differences analyses.