nature portfolio

Corresponding author(s): Shadi Albarquoni

Last updated by author(s): Jun 2, 2022

Reporting Summary

Nature Portfolio 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 Portfolio 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 for data collection as data was obtained from public datasets.
Data analysis	The code is made publicly available under https://github.com/albarqounilab/FedDis. The following packages are used: scikit_image==0.17.2 matplotlib==3.3.2 numpy==1.19.2 wandb==0.10.30 monai==0.6.0 plotly==5.1.0 PyYAML==5.4.1 scikit_learn==0.24.2 skimage==0.0 torch==1.9.0 Python >= 3.6

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 Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Most of the datasets used in this study are publicly available and can be downloaded after signing a Data Usage Agreement. The OASIS dataset is available at https://www.oasis-brains.org; the ADNI-S and ADNI-P datasets are available at http://adni.loni.usc.edu/data-samples/access-data/; the MSLUB dataset was available at http://lit.fe.uni-lj.si/tools.php?lang=eng; the MSISBI dataset is available at https://smart-stats-tools.org/lesion-challenge-2015; the WMH dataset is available at https://wmh.isi.uu.nl; and the BRATS 2018 dataset is available at https://www.med.upenn.edu/sbia/brats2018/data.html. For the prospective cohort, KRI, MSKRI, and GBKRI, upon reasonable request and signing of data transfer agreements and pending approval by our IRB and data protection officer, data can be shared.

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	All experimental design and data details are provided in the "Methods" section
Data exclusions	All experimental design and data details are provided in the "Methods" section
Replication	All experimental design and data details are provided in the "Methods" section
Randomization	Randomization was not performed as this is not applicable to our Image analysis work where we don't study treatment effects.
Blinding	Randomization was not performed as this is not applicable to our work as well

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	All demographic details are provided in Figure 1
Recruitment	For KRI, MSKRI, and GBKRI, all patients were part of in-house observational cohorts, some of which were prospective (MSKRI; with patient consent), while the others were retrospective (without patient consent).
Ethics oversight	For all patients, our local IRB approved the use of imaging data for research purposes after anonymization. As several patients were part of retrospective cohorts without explicit patient consent, these data cannot be shared as mandated by

our IRB. For the prospective cohort, upon reasonable request and signing of data transfer agreements and pending approval by our IRB and data protection officer, data can be shared.

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	Structural MRI (FLAIR)
Design specifications	All experimental design and data details are provided in the "Methods" section
Behavioral performance measures	not applicable

Acquisition

Imaging type(s)	Structural MRI
Field strength	3T
Sequence & imaging parameters	OASIS-3: Resolution: 5x1x1, TR/TE/T1: Diverse ADNI-S: Resolution: 5x0.9x0.9, TR/TE/T1: Diverse ADNI-P: Resolution: 5x0.9x0.9, TR/TE/T1: 9/0.09/2.5 KRI: Resolution: 1.5x0.9x0.9, TR/TE/T1: 10/0.14/2.75 MSLUB: Resolution: 0.8x0.47x0.47, TR/TE/T1: 5/0.392/1.8 MSISBI: Resolution: 2.2x0.82x0.82, TR/TE/T1: 11/0.068/2.8 MSKRI: Resolution: 1.5x0.9x0.9, TR/TE/T1: 10/0.14/2.75 GBKRI: Resolution: 1.5x0.9x0.9, TR/TE/T1: 5/0.395/1.8 BRATS: Resolution: 1x1x1, TR/TE/T1: Diverse WMH: Resolution: Diverse, TR/TE/T1: Diverse
Area of acquisition	whole-brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	All scans have been registered to the SRI24 atlas template space to ensure all data share the same volume size and orientation. Subsequently, the scans have been skull-stripped with ROBEX.
Normalization	The brain MRI volumes were linearly normalized to the [0,1] range.
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	No.
Volume censoring	No

Statistical modeling & inference

Model type and settings	deep learning methods were for processing the imaging data
Effect(s) tested	the ability to detect pathologies without providing annotations

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

Statistic type for inference (See Eklund et al. 2016)	pixel-wise reconctruction
Correction	No correction was needed

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	Deep auto-encoding models were used in this manuscript to learn the normative prior of healthy brain

Multivariate modeling and predictive analysis

images. This was used later on to reconstruct a pseudo healthy version of a given input data. The residual is then used to detect and segment anomalies (pathologies) in Brain MRI Imaging. DICE Score was used to evaluate the predicted segmented pathology.