Supplemental Online Content

Wen J, Fu CHY, Tosun D, et al; iSTAGING consortium; ADNI; BIOCARD; BLSA. Characterizing heterogeneity in neuroimaging, cognition, clinical symptoms, and genetics among patients with late-life depression. *JAMA Psychiatry*. Published online March 9, 2022. doi:10.1001/jamapsychiatry.2022.0020

eTable 7. Demographic and dimensional information of the ADNI, BLSA and BIOCARD longitudinal datasets

This supplemental material has been provided by the authors to give readers additional information about their work.

Multidimensional representations in late-life depression: convergence in neuroimaging, cognition, clinical symptomatology and genetics

- **Supplementary materials**

Junhao Wen, PhD[1,*], Cynthia H.Y. Fu, MD, PhD[2,3], Duygu Tosun, PhD[4], Yogasudha Veturi, PhD[5], Zhijian Yang, MS[1], Ahmed Abdulkadir, PhD[1], Elizabeth Mamourian, MS[1], Dhivya Srinivasan, MS[1], Ioanna Skampardoni, MS[1], Ashish Singh, MS[1], Hema Nawani, PhD[1], Jingxuan Bao, MS[6], Guray Erus, PhD[1], Haochang Shou, PhD[1,7], Mohamad Habes, PhD[8], Jimit Doshi, MS[1], Erdem Varol, PhD[9], R Scott Mackin, PhD[10], Aristeidis Sotiras, PhD[11], Yong Fan, PhD[1], Andrew J. Saykin, PhD[12], Yvette I. Sheline, MD, PhD[13], Li Shen, PhD[6], Marylyn D. Ritchie, PhD[5], David A. Wolk, MD, PhD[1,14], Marilyn Albert, PhD[15], Susan M. Resnick, PhD[16], Christos Davatzikos, PhD[1,*]

[1]Center for Biomedical Image Computing and Analytics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA
[2] University of East London, School of Psychology, London, UK
[3] Centre for Affective Disorders, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[4]Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA, USA
[5]Department of Genetics and Institute for Biomedical Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[6]Department of Biostatistics, Epidemiology and Informatics University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA
[7]Penn Statistics in Imaging and Visualization Center, Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA
[8]Glenn Biggs Institute for Alzheimer's & Neurodegenerative Diseases, University of Texas Health Science Center at San Antonio, San Antonio, USA
[9]Department of Statistics, Center for Theoretical Neuroscience, Zuckerman Institute, Columbia University, New York, USA
[10]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, CA, USA
[11]Department of Radiology and Institute for Informatics, Washington University School of Medicine, St. Louis, USA
[12]Radiology and Imaging Sciences, Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana Alzheimer's Disease Research Center and the Melvin and Bren Simon Cancer Center, Indiana University School of Medicine, Indianapolis
[13]Center for Neuromodulation in Depression and Stress, Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[14]Department of Neurology and Penn Memory Center, University of Pennsylvania, Philadelphia, USA
[15]Department of Neurology, Johns Hopkins University School of Medicine, USA
[16]Laboratory of Behavioral Neuroscience, National Institute on Aging, Baltimore, USA
[&] For the iSTAGING consortium, the ADNI, the BIOCARD, and the BLSA studies

[*]Corresponding authors:
Junhao Wen, PhD – junhao.wen89@gmail.com
Christos Davatzikos, PhD – Christos. Davatzikos@pennmedicine.upenn.edu
3700 Hamilton Walk, Philadelphia, PA 19104

Search terms: late-life depression; heterogeneity; semi-supervised clustering; dimensional representation

eMethods

eTable 5: Comparison of demographic, diagnostic, clinical, and cognitive variables between Dim1 and Dim2

eTable 6: Comparison of demographic, diagnostic, clinical, and cognitive variables between Dim1 and Dim2 in UKBB validation sample.

eTable 7: Demographic and dimensional information of the ADNI, BLSA and BIOCARD longitudinal datasets.

eMethod 1. Main datasets of the iSTAGING consortium used in this study.

This study investigates the neuroanatomical heterogeneity of LLD patients, including the UK Biobank (UKBB) (1), the UCSF, the Baltimore Longitudinal Study of Aging (BLSA) (2, 3), and The Biomarkers of Cognitive Decline Among Normal Individuals in the Johns Hopkins University (BIOCARD).

The UKBB dataset approximately recruits 500,000 UK adults sampled via population-based registries (http://www.ukbiobank.ac.uk). The UK Biobank received ethical approval from the National Research Ethics Service Committee North West–Haydock (reference 11/NW/0382). All participants provided informed consent and were aged from approximately 40 to 69 years of age at the time of enrollment (http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=200). Participants were recruited from across the United Kingdom, and initial enrollment was carried out from 2006 to 2010. Participants provided socio-demographic, cognitive, and medical data via questionnaires and physical assessments. Starting in 2014, a subset of the original sample later underwent magnetic-resonance brain imaging (MRI) (UK Biobank Brain Imaging Documentation, http://www.ukbiobank.ac.uk). The MRI data used in the current study were acquired between 2014 and 2019. To sum up, the image protocols of T1w MRI are as follows: a 3 T Siemens Skyra machine (MPRAGE) was used; image resolution: 1 x 1 x 1 mm and time to echo (TE): 2000 ms (4). More details can be found at http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf. The LLD inclusion criteria are:

- We first restricted the subjects to the age of 60 years and older (*N*=14664). Moreover, we excluded the following comorbidities based on the ICD10 and ICD9 information from UKBB (data field: 41202): 41 subjects with schizophrenia diagnosis (F2), four subjects with bipolar diagnosis (F31), 0 subjects with psychotic symptoms (F323, F333), 576 subjects with anxiety diagnosis (F41), seven subjets with Obsessive-compulsive disorder (F429), ten subjects with Posttraumatic stress disorder (F431), one subject with Huntington's disease (G10), two subjects with AD diagnosis (G30), 216 subjects with Epilepsy (G40), 50 subjects with Stroke (I64), 1214 subjects with Diabetes diagnosis, and 6342 subjects with a hypertension diagnosis (*N*=9182). Subsequently, LLD patients were selected based on data fields 20124 (probable recurrent major depression - moderate), and 20125 (probable recurrent major depression - severe), where LLD subjects were positive in either field 20124 or 20125 and control subjects were negative in both fields 20124 and

20125. This resulted in 425 LLD patients and 8757 healthy controls. We then matched 425 healthy controls by sex and age of the patients. Finally, we discarded the subjects that failed for the image preprocessing pipeline. This resulted in 397 healthy control (CN) and 402 LLD subjects. Moreover, the FA value of the participants is also provided by the UKBB (UKBB Application Number 35148).

The BLSA dataset is an ongoing longitudinal study led by the National Institute of Aging (NIA). The BLSA began in Baltimore, Maryland, the USA, in 1958, with the aim of characterizing the process of aging in general and healthy aging in particular. The BLSA acquired physiological measures, biomarkers, risk factors, blood samples and evaluated cognitive function in healthy individuals over the course of their lifetime. The neuroimaging sub-study of BLSA began in 1994, with annual or semi-annual imaging studies, cognitive testing, and clinical evaluations of a subset of 721 BLSA participants. Beginning in 1986, BLSA introduced an extensive neuropsychological assessment to subjects older than 60 years. BLSA participants underwent a T1w MRI using two different imaging protocols. At its early stage, a 1.5 T GE Signa machine (SPGR) was used: image resolution: 0.94 x 0.94 x 1.5 mm; flip angle: 45 degrees; time to echo (TE): 5 ms and repetition time (TR): 35 ms (3). More recently, a 3 T Philips machine (MPRAGE) was adopted: image resolution: 1 x 1 x 1.2 mm; flip angle: 8 degrees; time to echo (TE): 3.1 ms and repetition time (TR): 6.5 ms (5). We included the BLSA population based on the following criteria:

- All participants were restricted to be 60 years old or older. We first excluded the following comorbidities: 58 subjects with a diabetes diagnosis, 525 subjects with hypertension diagnosis, no subject with anxiety, bipolar and schizophrenia diagnoses. LDD patients (N=30) were selected with CES-Depression scores 16 and above and healthy control subjects (N=204) with CES-Depression scores six and below. Finally, we discarded the subjects that failed for the image preprocessing pipeline. This resulted in 27 CN and 29 LLD subjects that are age- and sex-matched.

The BIOCARD study was initiated by the Geriatric Branch of the National Institute of Mental Health in 1995. In order to identify biomarkers associated with the progression of cognitive impairment, BIOCARD recruited 354 cognitively healthy adults with a first-degree relative affected with AD. BIOCARD participants were initially admitted for a three-day-long evaluation

at the Clinical Center at the National Institutes of Health (NIH). At this time, a detailed physical examination, neuropsychological assessment, brain MRI, and blood and cerebrospinal fluid (CSF) samples were obtained. Subsequently, annual evaluations were conducted at the NIH until the study closed in 2005. The Johns Hopkins School of Medicine re-opened the BIOCARD study in 2009, continuing with annual clinical and cognitive follow-up for the original BIOCARD cohort. Amyloid PET imaging using the Pi-B tracer began in 2015, and Tau PET imaging began in 2017. In BIOCARD, Johns Hopkins University continued to administer all of the cognitive tests from the original protocol from NIH. The annual neuropsychological assessment covers every major cognitive domain (memory, executive function, language, visuospatial ability, attention, speed of processing, and psychomotor speed). This assessment includes TMT and CVLT cognitive tests of executive function and memory. The image protocols of T1w MRI are as follows: a 3 T Phillips machine (MPRAGE) was used; image resolution: 1 x 1 x 1.2 mm; image dimension: 256 x 256 x 170; flip angle: 8 degrees; and time to echo (TE): 3.1 ms and repetition time (TR): 6.75 ms.

- All participants were restricted to be 60 years old or older. Forty-nine subjects with a diabetes diagnosis, 152 subjects with a hypertension diagnosis. LLD subjects (*N*=6) are selected with Hamilton depression scores eight and above and healthy control subjects (*N*=58) with Hamilton depression scores seven and below. Finally, this resulted in 6 LLD patients and five age and sex-matched healthy controls after the quality check of image preprocessing pipelines.

The UCSF participants included community-dwelling adults age 65 and older. Inclusion criteria were: 1) Current DSM-IV diagnosis of major depressive disorder (MDD), unipolar type, without psychotic features, and six weeks minimum duration of a current depressive episode of moderate severity (24-item Hamilton depression score >20). 2) English speaking, male or female participants > 65 years of age. 3) Good general health, i.e., no evidence for uncontrolled medical illness. 4) No antidepressant use or psychotherapy within the past six weeks or electroconvulsive therapy within the past six months. 5) No recent (< 6 months) history of substance or alcohol abuse or dependence (DSM-IV criteria). 6) No cognitive-enhancing medications use. 7) No current diagnosis of Post-Traumatic Stress Disorder or other Axis I psychiatric disorders. 8) No significant neurological diseases (e.g., Parkinson's disease, epilepsy, cortical stroke, AD, traumatic brain injury) or dementia. 9) No history of surgical procedures affecting study outcomes. A subset of

Alzheimer's Disease Neuroimaging (ADNI) study participants with no current or past history of MDD or other psychiatric diagnosis and otherwise meet the same inclusion criteria as LLD participants were selected as healthy control reference for UCSF LLD cohort as the same neuroimaging and cognitive assessments were utilized in both studies. Diagnoses were made by a licensed clinical psychologist using DSMIV criteria. Participants with other Axis I disorders, significant current neurologic illness or diagnoses, or evidence of dementia (<25 on the Mini-Mental Status Exam) were excluded. Because the study was not a randomized control trial, there was no control condition. Eligible participants completed baseline assessment measures, including current suicidal ideation and measures of disability. Participants then participated in 12 weeks of individual problem-solving therapy (60 minutes per week). After therapy concluded, a follow-up assessment was conducted with duplicate measures for baseline comparison. The image protocols of T1w MPRAGE MRI are as follows: a 3 T Siemens Skyra scanner was used; image resolution: 1 x 1 x 1.2 mm; image dimension: 256 x 256 x 176; flip angle: 9 degrees; and time to echo (TE): 3 ms and repetition time (TR): 2300 ms. This study was funded by the National Institute of Mental Health (R01MH101472; PI: Tosun/Mackin). LLD was defined with the following criteria:

- Current diagnosis of unipolar Major Depressive Disorder without psychotic features, with a 24-item Hamilton Depression Rating Scale (HDRS) total score ≥19, and current depressive episode lasting at least six weeks. A subset of Alzheimer's Disease Neuroimaging (ADNI) study participants with no current or past history of MDD or other psychiatric diagnosis and otherwise meeting the same inclusion criteria as LLD participants were selected as a healthy control reference as the same neuroimaging and cognitive assessments were utilized in both ADNI and UCSF studies. This resulted in 65 LLD patients and corresponding 65 healthy controls. No participants have been diagnosed with other psychiatric disorders. More than 40% of UCSF LLD participants have cognitive impairment.

eMethod 2. Quality check steps for T1-weighted imaging preprocessing.

Raw 3D T1-weighted MRIs were first quality checked (QC) for motion, image artifacts, or restricted field-of-view. Another QC was performed as follows: First, the images were examined by manually evaluating for pipeline failures (e.g., poor brain extraction, tissue segmentation, and registration errors). Furthermore, a second-step automated procedure automatically flagged images based on outlying values of quantified metrics (i.e., ROI values), and those flagged images were re-evaluated.

eMethod 3. Image processing for UKBB diffusion tensor imaging data.

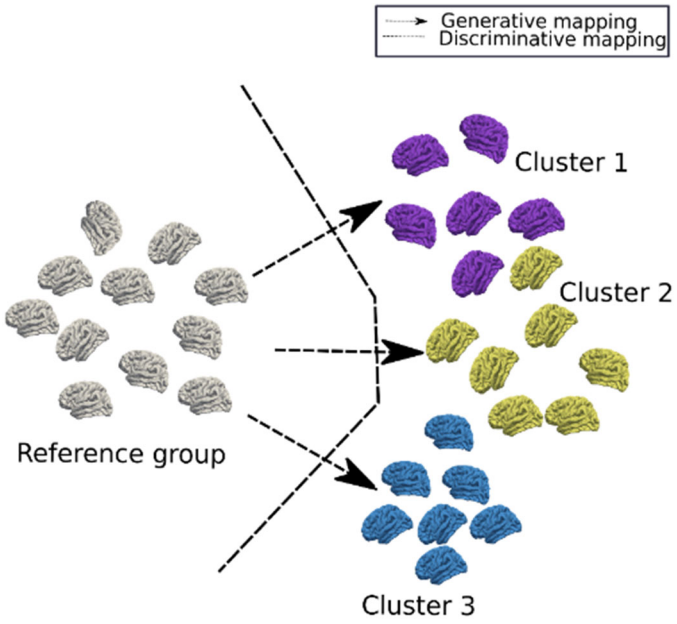
Details can be found at http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf. Briefly, diffusion MRI data were acquired with b-value=1000 s/mm^2 with 50 encoding directions. Moreover, b-value=0 images (b0) with opposite phase-encoding direction (anterior-posterior and posterior-anterior) were also obtained. Raw diffusion MRI data were first corrected for head motion and eddy current using *Eddy* tool (6). The pairs of b0 images were then fed into *Topup* (7) to estimate the b0 fieldmap and correct susceptibility-induced distortions. The corrected images were fit with the DTI model using *DTIFIT* tool (8) to generate the FA maps. Individual FA map was nonlinearly normalized in the MNI space using *FNIRT* (9).

eMethod 4. Inter-site image harmonization

The 119 MUSE ROIs (**Supplementary eTable 2**) from all four sites were harmonized before fitting to HYDRA. We applied an extensively validated statistical harmonization approach, i.e., ComBat-GAM (10), where variability in volumetric measures due to differences in site/cohort-specific imaging protocol differences was estimated based on variance observed within and across control groups while preserving normal variance due to age, sex, and intracranial volume (ICV) differences. The model was initially trained on the complete set of iSTAGING and was then applied to the LLD population to establish a harmonized pooled dataset.

eMethod 5. HYDRA

In contrast to the generative approach used in CHIMERA, HYDRA leverage a widely used discriminative method, i.e., support vector machines (SVM), to seek this "*1-to-k*" mapping. The novelty is that HYDRA extends multiple linear SVMs to the non-linear case in a piecewise fashion, thereby simultaneously serving for classification and clustering. Specifically, it constructs a convex polytope by combining the hyperplane from *k* linear SVMs, separating the CN group and the *k* subpopulation of the PT group. Intuitively, each face of the convex polytope can be regarded as a subtype, capturing a distinct disease effect. A schematic figure is shown below to illustrate the core idea behind HYDRA and semi-supervised clustering. One of the advantages of HYDRA and semi-supervised clustering is that it tends to avoid clustering patients according to disease-irrelevant confounds by directly clustering differences between controls and patients.



sFigure 1. Schematic diagram of HYDAR (Discriminative mapping). In general, HYDRA seeks a "*1-to-k*" mapping between a relatively homogeneous reference group (CN) and the subgroups of the heterogeneous target group (subtype), thereby teasing out clusters that are likely driven by distinct pathological trajectories instead of by unrelated heterogeneity, e.g., demographics or typical brain aging.

The convex polytope is estimated by sequentially solving each linear SVM as a sub-problem under the principle of the sample weighted SVM. In each iteration, only a subpopulation of the patient samples is taken into account during optimization. The optimization stops until the

sample weights get stable, i.e., the polytope was stably established. The objective of maximizing the polytope's margin can be summarized as:

$$\min_{\{\boldsymbol{w}_j,\boldsymbol{b}_j\}_{j=1}^{k}} \sum_{j=1}^{k} \frac{\|\boldsymbol{w}_j\|_2^2}{2} + \mu \sum_{i|y_i=+1}^{j} \frac{1}{k} \max\{0, 1 - \boldsymbol{w}_j^T \boldsymbol{X}_i^T - \boldsymbol{b}_j\} + \mu \sum_{i|y_i=-1}^{j} \boldsymbol{S}_{i,j} \max\{0, 1 + \boldsymbol{w}_j^T \boldsymbol{X}_i^T + \boldsymbol{b}_j\}$$

where $\boldsymbol{w}_j$ and $\boldsymbol{b}_j$ are the weight and bias for each hyperplane, respectively. μ is a penalty parameter on the training error, and $\boldsymbol{S}$ *is* the subtype membership matrix of dimension *N*x*k* deciding whether a patient sample *i* belongs to subtype *j*. The cluster membership is inferred as follows:

$$\boldsymbol{S_{i,j}} = \begin{cases} 1, & j = \underset{j}{\operatorname{argmax}}(\boldsymbol{w}_j^T \boldsymbol{X}^T + \boldsymbol{b}_j) \\ 0, & j \neq \underset{j}{\operatorname{argmax}}(\boldsymbol{w}_j^T \boldsymbol{X}^T + \boldsymbol{b}_j) \end{cases}$$

eMethod 6. The permutation test for the statistical significance of clustering solutions.

In the current study, the permutation test was performed with the following steps:

- We first split the CN group into the true CN (40%) and pseudo LLD (60%) groups. Run HYDRA with this data and obtain $ARI_{alternative\_i}$;
- We then take the same CN subjects from the previous step and choose the same number of true LLD patients. Run HYDRA with this data and obtain ARI_{null_i};
- Repeat the above two steps 250 times, and then compare the ARI_{null} and $ARI_{alternative}$ with a two-sample t-test.

If the P-value obtained is lower than the threshold (0.05), we conclude that the k subtypes are statistically significant.

eMethod 7. Applying the trained HYDRA model to test data.

After the model was trained on the LLD population, we saved the optimal polytope that best discriminated the CN group from the k dimensions. The polytope was positioned in the high-dimensional space by the weight w_i and bias b_i, where i indicates the i-th linear SVM.

The k sets of weight and bias were then applied to the validation UKBB samples, resulting in k SVM scores for each participant. We refer these scores to the expression score of each dimension. Mathematically, the expression score (E) can be written via distance from respective hyperplanes of SVMs as,

$$E_i = w_i \boldsymbol{X} + b_i$$

where $\boldsymbol{X}$ is the feature matrix (119 GM ROI volumes) of UKBB participants. Intuitively, a larger score represents that this specific dimension has been more expressed, and a smaller score means a more CN-like anatomy expression. Moreover, the hard-coded dimension membership (D) for each individual can be decided as follows:

$$D = \begin{cases} -1 & if\ E_i < 0, \forall i \in 1, 2, \dots k \\ \underset{i}{\operatorname{argmax}}(E_i) & otherwise \end{cases}$$

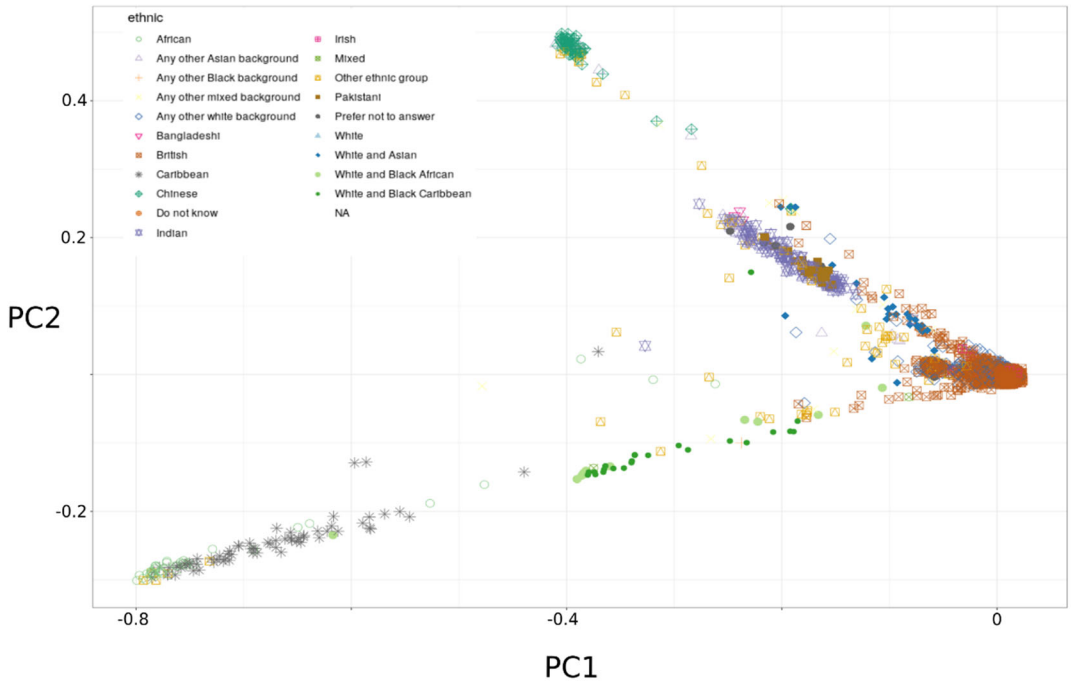
eMethod 8. Genetic data, heritability analysis, and GWAS in UKBB sample.

Raw genetic data (Version 3) was downloaded from the UKBB website (https://www.ukbiobank.ac.uk/enable-your-research/about-our-data/genetic-data) on July 2021. We used the imputed genetic data in the current study. Details regarding the imputation can be referred to in the original UKBB paper (11). The genotyped and imputed single nucleotide polymorphisms (SNPs) from UKBB (11) were preprocessed with a standard QC protocol as follows.

Genetic preprocessing protocol

First, we extracted and excluded related individuals (*N*=36394, 2nd-degree related individuals) in the complete UKBB sample (*N*=488377) via King software[1] relationship inference (12). We then removed duplicated variants (*N*=22488) from all 22 autosomal chromosomes. Subsequently, we merged and consolidated a UKBB imaging-genetic population to have both data modalities (*N*=21305). Thirteen participants were removed because of having unmatched sex identities between genetically identified sex and self-acknowledged sex. Lastly, we excluded i) subjects with more than 3% missing genotypes; ii) variants with minor allele frequency less than 0.01; iii) variants with larger than 3% missing genotyping rate; iv) variants that failed the Hardy-Weinberge test at $1*10^{-10}$. Finally, 20438 participants and 8430655 SNPs passed the QC procedures and were used for subsequent genome analyses. To adjust the population stratification (13), we derived the first 40 genetic principle components (PC) using the FlashPCA software[2]. sFigure 2 demonstrates the population structure in the current imaging-genetic population.

[1] https://www.chen.kingrelatedness.com/
[2] https://github.com/gabraham/flashpca



sFigure 2. Each point represents a UKBB individual (N = 20438) of the current imaging-genetic population and is placed according to their first and second principal component (PC). Colors and shapes indicate the self-reported ethnic background of each individual.

Genome-wide association studies (GWAS)

For genetic variants of autosomes, we performed logistic regression for the binary traits (CN as 1 for control and dimension label as 2 for case). We included the 40 PCs, age at imaging, genetic sex, and a deep learning-based ICV measure as covariates in the model. For the 799 UKBB participants of the LLD, 385 CN, 175 Dim1, and 214 Dim2 passed the QC and were included. The genome-wide P-value threshold was set as $5x10^{-8}$. After obtaining the GWAS summary statistics for each dimension, we input them to FUMA to define the genomic risk loci and independent significant SNPs. Specifically, FUMA first identified independent significant variants that passed the significant threshold and were independent of other significant variants at $r^2 < 0.6$. Subsequently, FUMA constructed LD blocks by tagging all variants that and a minor allele frequency of >= 0.0005 and were in LD at least one of the independent significant variants. This includes variants from the reference panel (the 1000 Genomes) that may not be overlapped with the variants in the current study. Finally, the LD blocks of independent significant variants were merged to be a single genomic locus if they are physically close (<250 kilobases based on the

closest boundary variants of LD blocks). All other parameters stayed as default on FUMA online platform.

Heritability estimates in the general population

We estimated the proportion of variance explained by all autosomal genetic variants for the expression score of Dim1 and Dim2 in the general population in UKBB using GCTA-GREML analysis[3]. The adjusted covariates included the 40 PCs, age at imaging, genetic sex, age-square, age-sex interaction, age-square-sex interaction, and a deep learning-based ICV measure. The expression scores were first normalized using inverse normal transformation (14), then adjusted in a linear model using the above-mentioned covariates, and residualized scores were fit into GCTA. The heritability estimates were tested in one-sided likelihood ratio tests.

[3] https://cnsgenomics.com/software/gcta/#GREMLinWGSorimputeddata

eMethod 9. Regression models to estimate the longitudinal changes of the two dimensions in ADNI, BLSA, and BIOCARD datasets.

Estimate of the betas/ rate of change (RC) of age via linear mixed-effects models

We estimated the betas or the rate of change (RC) of age using a linear mixed-effects model. Specifically, we included age as the variable of interest, study and ICV as fixed effects and the slope and intercept of each participant as random effects. Mathematically, it can be written as:

$$Y_{i,j} = \beta_0 + \beta_1 X_{i,j} + \gamma_{0i} + \gamma_{1i} X_{i,j} + \varepsilon_{i,j}$$

where $Y_{i,j}$ is the j-th measured response for subject i, and $X_{i,j}$ is a covariate matrix for the fixed effects (i.e., sex and ICV and age). β_0 and β_1 are shared by all subjects, and the errors $\varepsilon_{i,j}$ is identically distributed with a mean zero. The random-effects parameters γ_{0i} and γ_{1i} follow a bivariate distribution.

There are two parts to define the RC of age for each participant: i) the fixed-beta of age in β_1 and ii) the random-beta of age for each participant in γ_{1i}, i.e., RC = $\gamma_{1i\_age} + \beta_{1\_age}$.

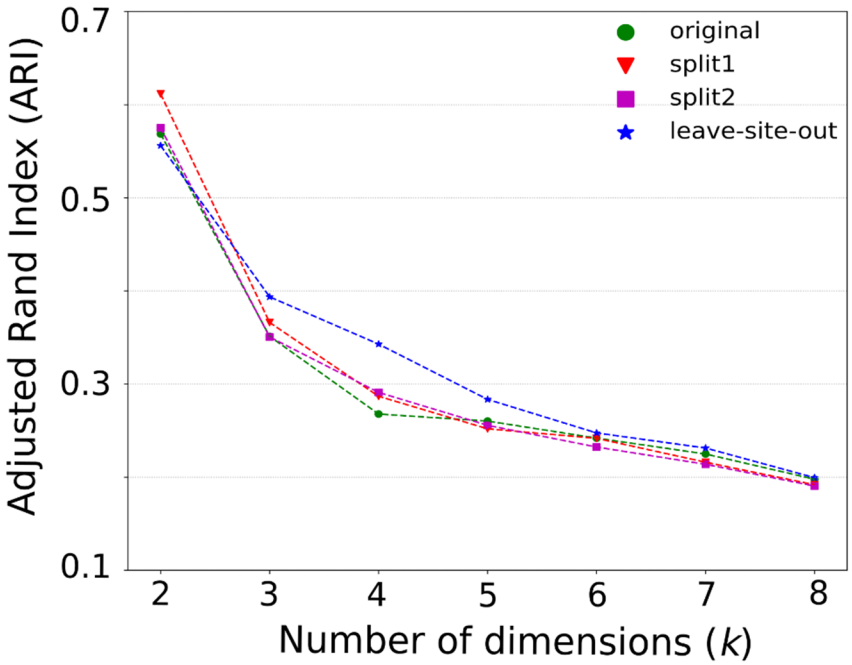
Group comparisons via linear models

After generating the RC of each participant, we performed a group comparison between two groups (e.g., Dim1 vs. Dim2) with a linear model, including age, sex, or ICV as covariates, to determine if the rate of change for the response variable ($Y_{i,j}$) statistically differ between groups.

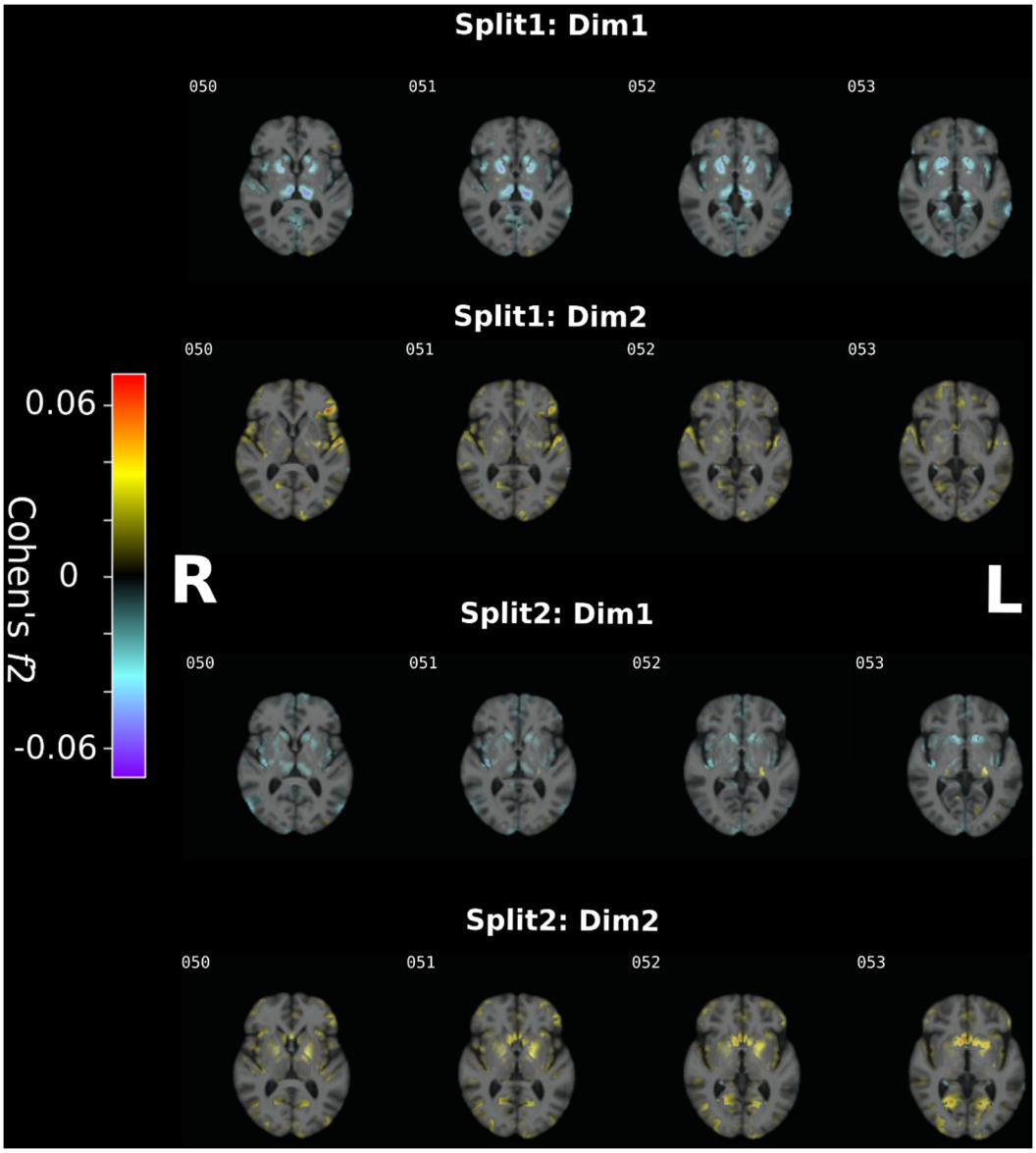
eMethod 10. Presentation of statistics.

We summarize the statistics used in the paper:

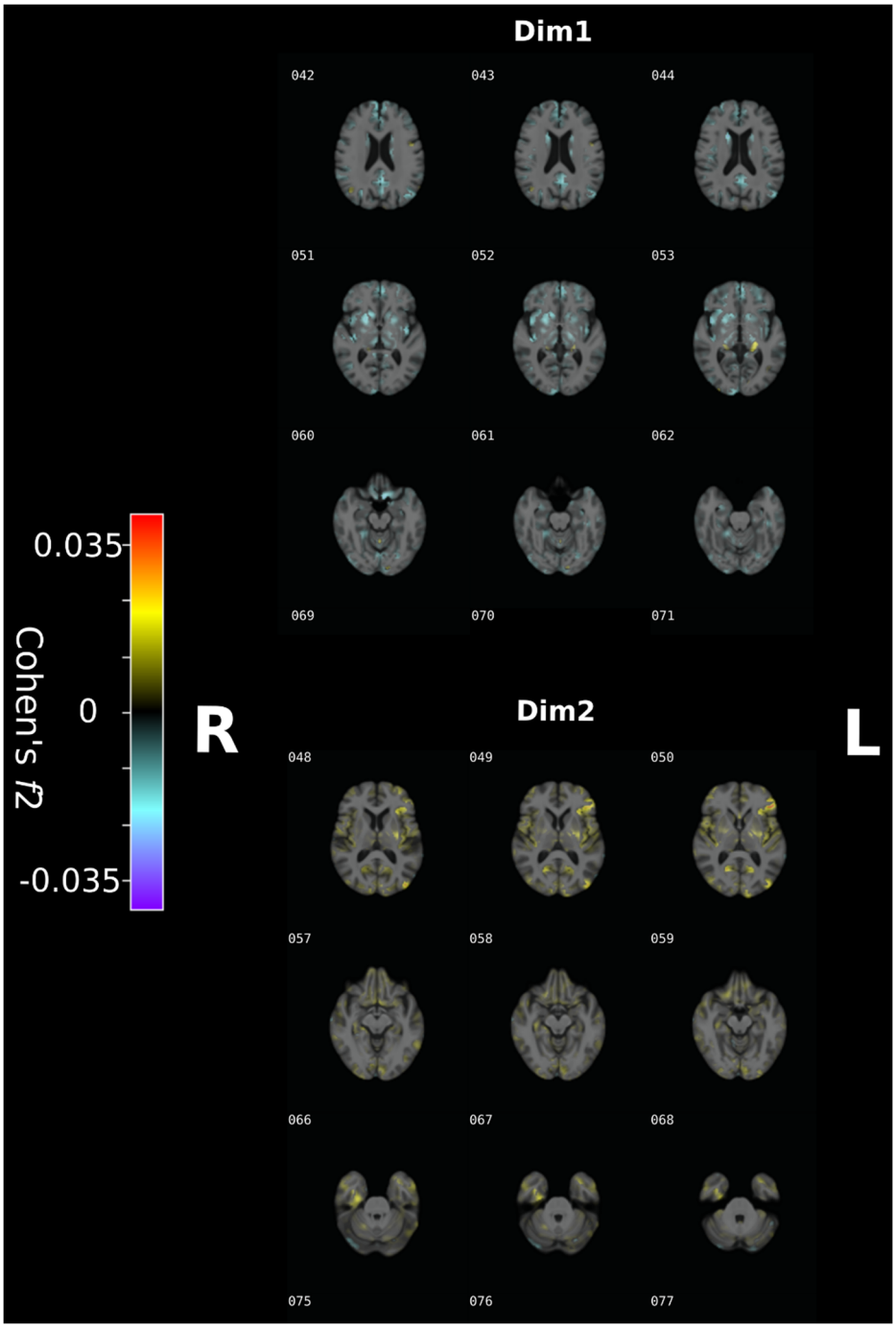
- For voxel-wise GM and WM tracts group comparisons, the multiple linear regression models were defined in the main text. We presented the results as an effect size (Cohen's f2) map for voxels/tracts that passed the P-value threshold (0.05) after the correction for multiple comparisons (Benjamini-Hochberg procedure).
- For group comparisons for demographic, cognitive, and clinical variables, each variable was examined separately between the two dimensions. Mann–Whitney–Wilcoxon test was used for continuous variables (e.g., age) and the Chi-Square test of independence for categorical variables (e.g., sex). Moreover, a global effect size (i.e., Cohen's *d*) was also reported for continuous variables. The threshold of the P-value was set to be 0.05
- For the heritability estimation, we reported the heritability coefficient (*h2*) and also the P-value. For GWAS analysis, we set the P-value threshold to be $5x10^{-8}$, which is a strict threshold in genome-wide analyses.
- For longitudinal analysis, Supplementary eMethod 9 detailed the regression models. The P-value threshold was set to be 0.05.



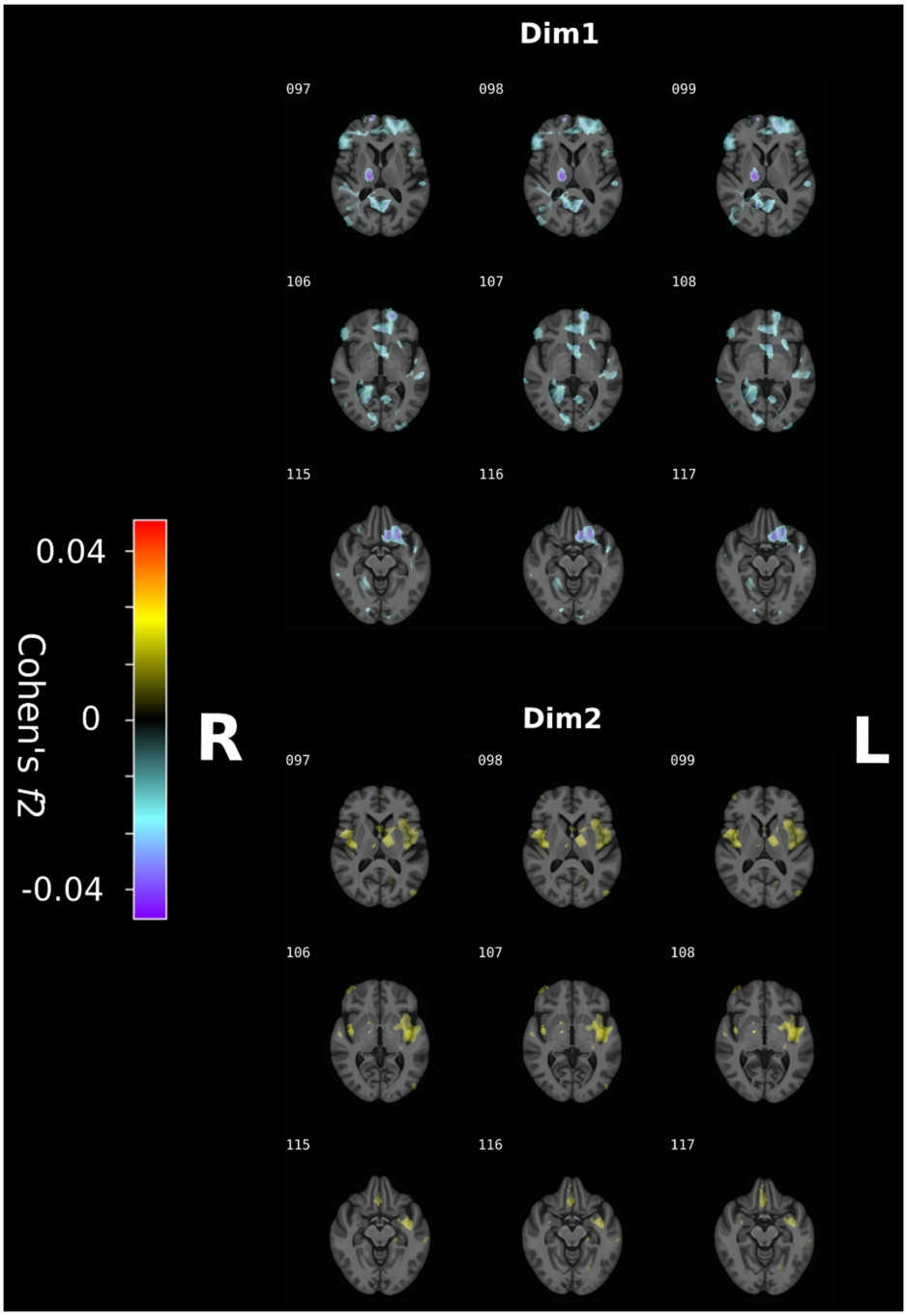
eFigure 1: Cross-validated clustering stability of the number of dimensions: Adjusted Rand Index (ARI) vs. the number of dimensions (k) indicating high reproducibility for $k = 2$ for all approaches. The green dot, red triangle, purple square, and blue star represent the original experiments, split-sample experiments for the first and second split, and the leave-site-out (training on UKBB) experiments, respectively.



eFigure 2: For split-sample analyses, effect size maps of GM patterns identified in Dim1 and Dim2 compared to controls, respectively. The two neuroanatomical dimensions show distinct GM abnormalities and highly reproducible between the two splits. Warm color denotes brain atrophy (i.e., CN > Dim) and cool color represents larger tissue volume (i.e., Dim > CN). Both directions are shown for each dimension. For references, Cohen's $f2$ of ≥ 0.02, ≥ 0.15, and ≥ 0.35 signify small, moderate, and large effect sizes, respectively. L: left; R: right. The effect size map is shown in a radiological fashion, i.e., the brain's left shown to the right of the display.

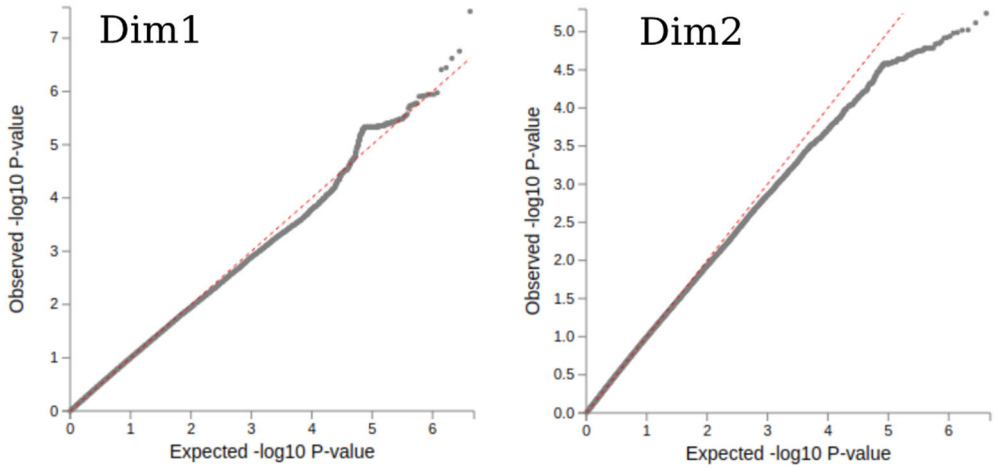


eFigure 3: For leave-site-out analyses (training only on UKBB MUSE ROIs), effect size maps of GM patterns identified in Dim1 and Dim2 compared to CN, respectively. The two neuroanatomical dimensions show distinct GM abnormalities. Warm color denotes brain atrophy (i.e., CN > Dim) and cool color represents larger tissue volume (i.e., Dim > CN). Both directions are shown for each dimension. For references, Cohen's $f2$ of ≥ 0.02, ≥ 0.15, and ≥ 0.35 signify small, moderate, and large effect sizes, respectively. L: left; R: right. The effect size map is shown in a radiological fashion, i.e., the brain's left shown to the right of the display.

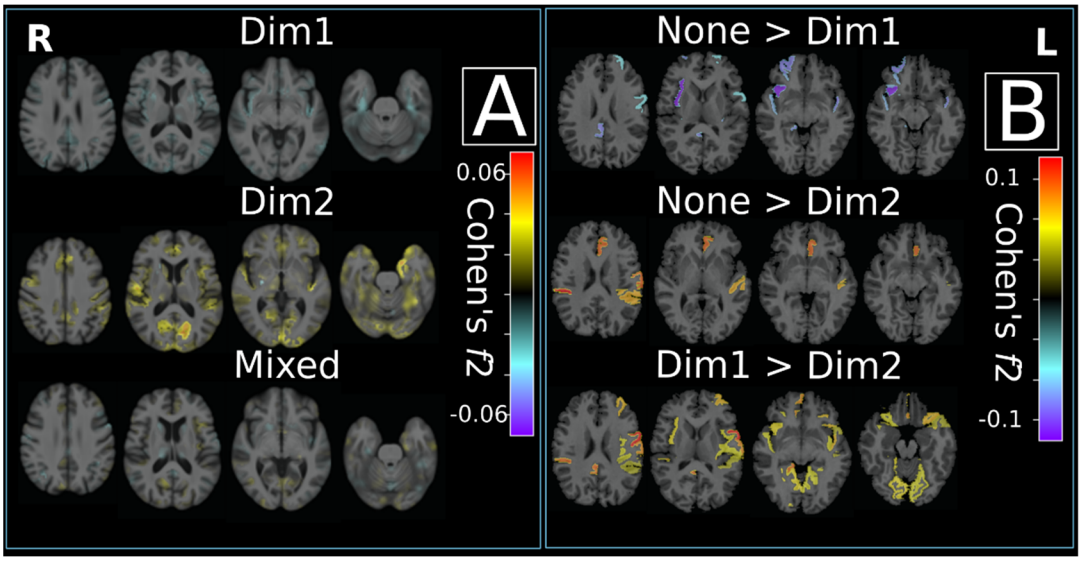


eFigure 4: The two dimensions derived from the population, including comorbidities from UKBB. This ended up with 532 LLD patients (501 in the original population with the inclusion/exclusion criteria). Effect size maps of GM patterns identified in Dim1 and Dim2 compared to CN, respectively. The two neuroanatomical dimensions show distinct GM abnormalities. Warm color denotes brain atrophy (i.e., CN > Dim) and cool color represents larger tissue volume (i.e., Dim > CN). Both directions are shown for each dimension. For references, Cohen's $f2$ of ≥ 0.02, ≥ 0.15,

and ≥ 0.35 signify small, moderate, and large effect sizes, respectively. L: left; R: right. The effect size map is shown in a radiological fashion, i.e., the brain's left shown to the right of the display.



eFigure 5: The quantile-quantile (Q-Q) plots for GWAS in Dim1 and Dim2 compared to CN, respectively. Q-Q plots take the actual GWAS P-values, sort them in ascending order, and then plot them versus quantiles calculated from a theoretical uniform distribution. The Q-Q plots show that the observed statistics are underpowered by the small sample sizes.



eFigure 6: The two dimensions and longitudinal trajectories. **A**) The two neuroanatomical dimensions in ADNI, BLSA, and BIOCARD baseline images show distinct grey matter abnormalities. Warmer color denotes brain atrophy (i.e., None > Dim), and cooler color represents larger tissue volume (i.e., Dim > None). Both directions are shown for each dimension. Cohen's *f*2 of ≥ 0.02, ≥ 0.15, and ≥ 0.35 signify small, moderate, and large effect sizes, respectively. L: left; R: right. **B**) The rate of change (RC) shows that Dim1's brain volume decreases with time more rapidly than Dim2. Only subjects for which MRI data were available at least for 6-time points were included for this analysis.

eTable 1. Demographic, diagnostic, cognitive, clinical, and imaging characteristics of study cohorts.

Age is shown with mean and its range. Education, and other continuous variables are shown with mean and standard deviation. Sex is displayed with the female and its percentage. Mann–Whitney–Wilcoxon test was used for continuous variables (e.g., age) and the Chi-Square test of independence for categorical variables (e.g., sex). CN: healthy control; RAVENS: gray matter voxel-wise map from T1-weighted MRI; MUSE: 119 GM regions of interests (ROIs) from T1-weighted MRI; FA: fractional anisotropy from diffusion MRI. --: not applicable; Y: yes; N: no; P: P-value. For additional cognitive scores per site, please refer to supplementary eTable 5.

	Harmonized (N=996)			UKBB (N=799)			BLSA (N=56)			BIOCARD (N=11)			UCSF (N=130)		
	CN	LLD	P	CN	LLD	P	CN	LLD	P	CN	LLD	P	CN	LLD	P
N	495	501		397	402		27	29		6	5		65	65	
Age [min/max]	66.26 [60, 91.47]	67.33 [60, 91]	0.34	65.28 [60, 76]	66.18 [60, 77]	0.32	67.70 [60, 84]	73.68 [60, 87]	0.02	63.81 [61, 70]	66.19 [62, 73]	0.12	71.84 [64,91.4]	71.67 [65, 91]	0.83
Sex/ female, %	333/ 67%	332/ 66%	0.78	264/ 66%	265/ 66%	0.92	21/ 77%	20/ 69%	0.65	5/ 83%	4/ 80%	1.00	43/ 66%	43/ 66%	1.00
Education/years	14.76 ± 2.68	14.87 ± 2.62	0.55	14.28 ± 2.50	14.26 ± 2.48	0.32	18.00 ± 1.41	17.69 ± 1.74	0.38	16.82 ± 2.39	16.86 ± 2.42	0.46	16.77 ± 2.62	16.40 ± 1.66	0.39
Systole	135.03 ± 16.83	134.75 ± 16.56	0.52	135.03 ± 16.83	134.75 ± 16.56	0.52	--	--	--	--	--	--	--	--	--
Diastole	75.59 ± 9.24	79.05 ± 9.15	0.45	75.59 ± 9.24	79.05 ± 9.15	0.45	--	--	--	--	--	--	--	--	--
Age of onset	--	34.62 ± 15.70	--	--	35.00 ± 14.34	--	--	--	--	--	--	--	--	32.68 ± 21.44	--
TMT-A	219.45 ± 74.82	227.76 ± 76.71	0.80	219.45 ± 74.82	227.76 ± 76.71	0.80	--	--	--	--	--	--	--	--	--
TMT-B	532.09 ± 251.43	549.80 ± 232.33	0.45	532.09 ± 251.43	549.80 ± 232.33	0.45	--	--	--	--	--	--	--	--	--
Imaging data availability															
RAVENS	Y		--	Y		--	Y		--	Y		--	Y		--
MUSE	Y		--	Y		--	Y		--	Y		--	Y		--
FA	Y		--	Y		--	N		--	N		--	N		--

eTable 2: Clustering features for HYDRA, including 119 MUSE GM ROIs.
L: Left hemisphere; R: Right hemisphere.

Precentral gyrus (R)	Occipital fusiform gyrus (R)	Anterior insula (L)
Precentral gyrus (L)	Planum temporale (R)	Anterior orbital gyrus (R)
Accumbens area (R)	Cerebellar vermal lobules I-V	Anterior orbital gyrus (L)
Accumbens area (L)	Cerebellar vermal lobules VI-VII	Angular gyrus (R)
Amygdala (R)	Cerebellar vermal lobules VIII-X	Angular gyrus (L)
Amygdala (L)	Basal forebrain (R)	Calcarine cortex (R)
Occipital pole (L)	Basal forebrain (L)	Calcarine cortex (L)
Caudate (R)	Middle temporal gyrus (L)	Central operculum (R)
Caudate (L)	Occipital pole (R)	Central operculum (L)
Cerebellum exterior (R)	Planum temporale (L)	Cuneus (R)
Cerebellum exterior (L)	Parietal operculum (L)	Cuneus (L)
Planum polare (L)	Postcentral gyrus (R)	Entorhinal area (R)
Middle temporal gyrus (R)	Postcentral gyrus (L)	Entorhinal area (L)
Hippocampus (R)	Posterior orbital gyrus (R)	Frontal operculum (R)
Hippocampus (L)	Temporal pole (R)	Frontal operculum (L)
Precentral gyrus medial segment (R)	Temporal pole (L)	Frontal pole (R)
Precentral gyrus medial segment (L)	Triangular part of the inferior frontal gyrus (R)	Frontal pole (L)
Superior frontal gyrus medial segment (R)	Triangular part of the inferior frontal gyrus (L)	Fusiform gyrus (R)
Superior frontal gyrus medial segment (L)	Transverse temporal gyrus (R)	Fusiform gyrus (L)
Pallidum (R)	Superior frontal gyrus medial segment (L)	Gyrus rectus (R)
Pallidum (L)	Planum polare (R)	Gyrus rectus (L)
Putamen (R)	Transverse temporal gyrus (L)	Inferior occipital gyrus (R)
Putamen (L)	Anterior cingulate gyrus (R)	Inferior occipital gyrus (L)
Thalamus proper (R)	Anterior cingulate gyrus (L)	Inferior temporal gyrus (R)
Thalamus proper (L)	Anterior insula (R)	Inferior temporal gyrus (L)
Lingual gyrus (R)	Occipital fusiform gyrus (L)	Subcallosal area (R)
Lingual gyrus (L)	Opercular part of inferior frontal gyrus (R)	Subcallosal area (L)
Lateral orbital gyrus (R)	Opercular part of inferior frontal gyrus (L)	Superior frontal gyrus (R)
Lateral orbital gyrus (L)	Orbital part of inferior frontal gyrus (R)	Superior frontal gyrus (L)
Middle cingulate gyrus (R)	Orbital part of inferior frontal gyrus (L)	Supplementary motor cortex (R)
Middle cingulate gyrus (L)	Posterior cingulate gyrus (R)	Supplementary motor cortex (L)
Medial frontal cortex (R)	Posterior cingulate gyrus (L)	Supramarginal gyrus (R)
Medial frontal cortex (L)	Precuneus (R)	Supramarginal gyrus (L)
Middle frontal gyrus (R)	Precuneus (L)	Superior occipital gyrus (R)
Middle frontal gyrus (L)	Parahippocampal gyrus (R)	Superior occipital gyrus (L)
Middle occipital gyrus (R)	Parahippocampal gyrus (L)	Superior parietal lobule (R)
Middle occipital gyrus (L)	Posterior insula (R)	Superior parietal lobule (L)
Medial orbital gyrus (R)	Posterior insula (L)	Superior temporal gyrus (R)
Medial orbital gyrus (L)	Parietal operculum (R)	Superior temporal gyrus (L)
Superior frontal gyrus medial segment (R)	Posterior orbital gyrus (L)	

eTable 3: The membership distribution across sites for different choices of number of dimensions.

Dimensions	Harmonized	UKBB	BLSA	BIOCARD	UCSF
k = 2	227	180	14	4	29
	274	222	15	1	36
k = 3	106	87	5	1	13
	221	177	12	1	31
	174	138	12	3	21
k = 4	113	91	3	3	16
	165	127	10	1	27
	125	103	9	1	12
	98	81	7	0	10
k = 5	81	69	4	1	7
	78	69	3	0	6
	61	49	3	2	7
	162	122	13	1	26
	119	93	6	1	19
k = 6	72	59	4	2	7
	126	96	7	0	23
	51	45	1	1	4
	79	68	1	1	9
	130	97	11	1	21
	43	37	5	0	1
k = 7	63	57	3	1	2
	126	92	12	1	21
	54	50	0	1	3
	47	37	3	2	5
	47	36	4	0	7
	99	77	6	0	16
	65	53	1	0	11
k = 8	71	54	2	1	14
	63	55	1	1	6
	85	66	10	0	9
	46	40	2	1	3
	50	40	4	1	5
	85	68	4	0	13
	60	51	0	0	9
	41	28	6	1	6

eTable 4: WM analyses for group comparison: Controls (CN) vs Dim1, CN vs Dim2 are all shown for 48 tracts from the JHU ICBM-DTI-81 WM label atlas. Multiple comparison corrected P-value/Cohen's *f2* are shown in the table. Note that Cohen's f2 ≥ 0.02, ≥ 0.15, and ≥ 0.35 signify small, moderate, and large effect sizes, respectively. * signifies significance. Positive Cohen's *f*2 indicates CN a larger fractional anisotropy (FA).

Tract	CN vs Dim1 (Pvalue/Cohen's *f*2)	CN vs Dim2(Pvalue/Cohen's *f*2)
Middle Cerebellar Peduncle (MCP)	0.746/0	**0.001*/0.05**
Pontine Crossing Tract (PCT)	0.395/0.01	**0.016*/0.02**
Genu of Corpus Callosum (GCC)	0.395/0.01	**0.001*/0.03**
Body of Corpus Callosum (BCC)	0.98/0	**0.026*/0.01**
Splenium of Corpus Callosum (SCC)	0.777/0	0.084/0.01
Fornix (Fx)	0.777/0	0.629/0
Left Corticospinal Tract (CST.L)	0.727/0	**0.026*/0.01**
Right Corticospinal Tract (CST.R)	0.306/0.01	**0.01*/0.02**
Right Medial Lemniscus (ML.R)	0.513/0.01	**0.006*/0.02**
Left Medial Lemniscus (ML.L)	0.98/0	**0.026*/0.01**
Right Inferior Cerebellar Peduncle (ICP.R)	0.934/0	0.232/0
Left Inferior Cerebellar Peduncle (ICP.L)	0.843/0	0.826/0
Right Superior Cerebellar Peduncle (SCP.R)	0.093/0.02	**0.017*/0.02**
Left Superior Cerebellar Peduncle (SCP.L)	0.363/0.01	**0.019*/0.01**
Right Cerebral Peduncle (CP.R)	0.809/0	**0.01*/0.02**
Left Cerebral Peduncle (CP.L)	0.777/0	**0.026*/0.01**
Left Anterior Limb of Internal Capsule (ALIC.L)	0.934/0	**0.015*/0.02**
Right Anterior Limb of Internal Capsule (ALIC.R)	0.98/0	**0.019*/0.01**
Left Posterior Limb of Internal Capsule (PLIC.L)	0.897/0	**0.027*/0.01**
Rgith Posterior Limb of Internal Capsule (PLIC.R)	0.843/0	**0.008*/0.02**
Left Retrolenticular Part of Internal Capsule (RLIC.L)	0.777/0	0.05/0.01
Right Retrolenticular Part of Internal Capsule (RLIC.R)	0.864/0	0.067/0.01
Left Anterior Corona Radiata (ACR.L)	0.98/0	0.107/0.01
Right Anterior Corona Radiata (ACR.R)	0.843/0	0.24/0
Left Superior Corona Radiata (SCR.L)	0.777/0	**0.019*/0.01**
Left Posterior Corona Radiata (PCR.L)	0.843/0	**0.021*/0.01**
Right Superior Corona Radiata (SCR.R)	0.843/0	**0.021*/0.01**
Right Posterior Corona Radiata (PCR.R)	0.934/0	**0.024*/0.01**
Left Posterior Thalamic Radiation (PTR.L)	0.98/0	**0.002*/0.03**
Right Posterior Thalamic Radiation (PTR.R)	0.98/0	**0.011*/0.02**
Right Sagittal Stratum (SS.R)	0.843/0	**0.017*/0.02**
Right External Capsule (EC.R)	0.843/0	**0.044*/0.01**
Left Sagittal Stratum (SS.L)	0.939/0	**0.009*/0.02**
Left External Capsule (EC.L)	0.843/0	**0.03*/0.01**
Left Cingulum Cingulate Gyrus (CGC.L)	0.809/0	**0.002*/0.03**
Left Cingulum Hippocampus (CGH.L)	0.777/0	0.069/0.01
Right Cingulum Cingulate Gyrus (CGC.R)	0.777/0	**0.008*/0.02**
Right Cingulum Hippocampus (CGH.R)	0.864/0	0.394/0
Left Fornix (Cres) / Stria Terminalis (Fx.ST.L)	0.98/0	0.46/0
Left Superior Longitudinal Fasciculus (SLF.L)	0.98/0	**0.045*/0.01**
Left Superior Fronto-Occipital Fasciculus (SFO.L)	0.513/0.01	0.114/0.01
Right Fornix (Cres) / Stria Terminalis (Fx.ST.R)	0.897/0	0.121/0.01
Right Superior Longitudinal Fasciculus (SLF.R)	0.972/0	**0.03*/0.01**
Right Superior Fronto-Occipital Fasciculus (SFO.R)	0.578/0	0.069/0.01
Left Uncinate Fasciculus (UNC.L)	0.093/0.02	0.61/0
Right Uncinate Fasciculus (UNC.R)	0.395/0.01	0.969/0
Left Tapatum (TAP.L)	0.934/0	**0.037*/0.01**
Right Tapatum (TAP.R)	0.478/0.01	0.174/0

eTable 5: Comparison of demographic, diagnostic, clinical, and cognitive variables between Dim1 and Dim2.

Mann–Whitney–Wilcoxon test was used for continuous variables, and the Chi-Square test of independence was performed for categorical variables. P-value and Cohen's *d* are both reported. For references, Cohen's $d \geq 0.2$, ≥ 0.5, and ≥ 0.8 signify small, medium, and large effect sizes, respectively. * signifies significance, and the results are bolded. Positive Cohen's *d* indicates that the value of Dim1 is greater than the value of Dim2, and vice versa. Antidepressant taken is a categorical variable that indicates if the participant has taken antidepressant medications, including sertraline, citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, amitriptyline, amoxapine, clomipramine, desipramine, doxepin, imipramine, nortriptyline, protriptyline, trimipramine, and maprotiline. DSF: digital span forward; DSB: digital span backward; DSST: digit symbol substitution test; FIS: fluid intelligence score; NA: not applicable. For DSST in UKBB, the scores were derived from a questionnaire: "About how old were you the FIRST time you had a period of two weeks like this? (Whether or not you received any help for it)".

Variable	P-value	Cohen's *d*	Variable	P-value	Cohen's *d*
Age	0.08	0.12	Matrix Pattern (UKBB)	0.19	0.19
Sex	0.83	NA	Tower Rearranging (UKBB)	0.89	0.01
Education	0.46	0.02	HAMD24Total (UCSF)	0.20	0.29
Systole	0.93	-0.02	HAMD17Total (UCSF)	0.34	0.29
Diastole	0.08	-0.15	AVLT STPR (UCSF)	0.46	-0.05
TMT-A (UKBB)	0.78	-0.15	GDS Total (UCSF)	0.24	-0.05
TMT-B (UKBB)	0.56	-0.06	**PHQ9 Total (UCSF)**	**0.04***	**-0.45**
DSF (UCSF)	0.25	0.27	CCI Total20 (UCSF)	0.25	-0.28
DSB (UCSF)	0.22	0.02	BLWHOTOT (UCSF)	0.28	-0.21
Pairs Matching (UKBB)	**0.003***	**-0.28**	AVLT Recognition (UCSF)	0.06	-0.04
DSST	0.46	0.16	BNT Raw (UCSF)	0.10	0.40
AVLT LOT (UCSF)	0.4	-0.06	AVLT Trial1 (UCSF)	0.15	0.41
Pros Mem (UKBB)	0.68	0.20	AVLT TrialB (UCSF)	0.37	-0.003
Reaction Time (UKBB)	0.98	0.003	AVLT Trial6 (UCSF)	0.36	-0.15
FIS (UKBB)	**0.008***	**0.25**	Weight change during depression	0.13	NA
Thoughts of death during depression	0.98	NA	Disease duration (UKBB)	0.43	0.07
Age of onset (UKBB)	0.13	-0.16	Depression longest period (weeks) (UKBB)	0.12	0.17
Number of episodes (UKBB)	0.18	0.16	Seen a psychiatrist (UKBB)	0.89	NA
Non-medication activity treatment (e.g., psychotherapy) (UKBB)	0.25	NA	Difficulty in concentration (UKBB)	0.45	NA
Impact on daily life during depression (UKBB)	0.86	NA	Frequency of depressed days during depression (UKBB)	0.32	NA
Fraction of time affected by depression (day) (UKBB)	0.97	NA	AVLT Delayed (UCSF)	0.18	-0.19
Feeling of worthlessness during depression (UKBB)	0.18	NA	Feeling of tiredness during depression (UKBB)	0.62	NA

Antidepressant taken (UKBB)	0.45	NA	Site	0.45	NA

eTable 6: Comparison of demographic, diagnostic, clinical, and cognitive variables between Dim1 and Dim2 in UKBB validation sample.
Mann–Whitney–Wilcoxon test was used for continuous variables, and the Chi-Square test of independence was performed for categorical variables. P-value and Cohen's *d* are both reported. For references, Cohen's $d \geq 0.2$, ≥ 0.5, and ≥ 0.8 signify small, medium, and large effect sizes, respectively. * signifies significance, and the results are bolded. Positive Cohen's *d* indicates that the value of Dim1 is greater than the value of Dim2, and vice versa. FIS: fluid intelligence score; NA: not applicable.

Variable	None (*N*=3500)	Dim1 (*N*=2269)	Dim2 (*N*=3786)	Mixed (*N*=2963)	P-value (Dim1 vs Dim2)	Cohen's *d* (Dim1 vs Dim2)
Age	67.14±4.52	67.15±4.43	67.27±4.55	67.20±4.48	0.37	-0.04
Sex/female, %	1762/50%	1076/47%	1803/47%	1482/50%	0.90	NA
TMT-A	226.74±91.75	222.24±82.08	235.99±90.24	229.27±98.13	**<0.0001***	**-0.16**
TMT-B	571.42±277.45	550.12±260.21	591.66±302.63	552.89±268.00	**<0.0001***	**-0.14**
DSST	18.09±4.97	18.21±4.98	17.64±5.06	18.38±5.04	**<0.001***	**0.11**
Pairs Matching	0.41±0.96	0.34±0.81	0.46±1.02	0.40±0.91	**<0.0001***	**-0.13**
Pros Mem	1063.78±717.14	1052.56±789.14	1047.31±776.78	1033.21±696.45	0.28	0.006
Reaction Time	552.35±100.41	555.48±103.06	554.89±104.89	554.27±103.96	0.49	0.005
FIS	6.75±2.00	7.01±2.09	6.43±2.02	6.78±2.04	**<0.0001***	**0.28**
Matrix Pattern	7.81±2.06	8.01±2.01	7.63±2.09	7.93±2.09	**<0.0001***	**0.19**
Tower Rearranging	9.72±3.16	9.86±3.18	9.52±3.17	9.80±3.24	**<0.01***	**0.11**

eTable 7: Demographic and dimensional information of the ADNI, BLSA and BIOCARD longitudinal datasets.
Mean follow-up years are presented with mean and first and third quantiles.

Study	*N* (participants/scans)	None (*N*=410)	Dim1 (*N*=301)	Dim2 (*N*=390)	Mixed (*N*=330)	Mean follow-up years	Sex (F/%)	Age
ADNI	612/3466	194	132	143	143	1.15 [0.28, 4.07]	329/54%	73.39±6.01
BLSA	638/3181	185	116	185	152	4.00 [1.00, 7.00]	331/52%	72.09±8.01
BIOCARD	181/577	31	53	62	35	2.20 [0, 10.67]	105/58%	66.01±5.07

References

1. Miller KL, Alfaro-Almagro F, Bangerter NK, et al.: Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nature Neuroscience 2016; 19:1523–1536

2. Resnick SM, Pham DL, Kraut MA, et al.: Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. J Neurosci 2003; 23:3295–3301

3. Resnick SM, Goldszal AF, Davatzikos C, et al.: One-year age changes in MRI brain volumes in older adults. Cereb Cortex 2000; 10:464–472

4. Alfaro-Almagro F, Jenkinson M, Bangerter NK, et al.: Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. Neuroimage 2018; 166

5. Habes M, Sotiras A, Erus G, et al.: White matter lesions: Spatial heterogeneity, links to risk factors, cognition, genetics, and atrophy. Neurology 2018; 91:e964–e975

6. Andersson JLR, Sotiropoulos SN: An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. Neuroimage 2016; 125:1063–1078

7. Andersson JLR, Skare S, Ashburner J: How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. NeuroImage 2003; 20:870–888

8. Basser PJ, Mattiello J, Lebihan D: Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo. Journal of Magnetic Resonance, Series B 1994; 103:247–254

9. Andersson JLR, Jenkinson M, Smith S: Non-linear registration aka Spatial normalisation. FMRIB Technial Report TR07JA2 2010;

10. Pomponio R, Erus G, Habes M, et al.: Harmonization of large multi-site imaging datasets: Application to 10,232 MRIs for the analysis of imaging patterns of structural brain change throughout the lifespan [Internet]. Bioinformatics, 2019[cited 2019 Oct 10] Available from: http://biorxiv.org/lookup/doi/10.1101/784363

11. The UK Biobank resource with deep phenotyping and genomic data | Nature [Internet][cited 2021 Sep 10] Available from: https://www.nature.com/articles/s41586-018-0579-z#Sec1

12. Manichaikul A, Mychaleckyj JC, Rich SS, et al.: Robust relationship inference in genome-wide association studies. Bioinformatics 2010; 26:2867–2873

13. Price AL, Zaitlen NA, Reich D, et al.: New approaches to population stratification in genome-wide association studies. Nat Rev Genet 2010; 11:459–463

14. Yang J, Loos RJF, Powell JE, et al.: FTO genotype is associated with phenotypic variability of body mass index. Nature 2012; 490:267–272