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

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], Scott R 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, 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
Word counts: 3000
Date of revision: Nov 30, 2021

Key points

Question: Is late-life depression (LLD) associated with one or multiple structural neuroimaging patterns?

Findings: Two dimensions best represented LLD neuroanatomical heterogeneity. Dimension 1 was associated with preserved brain structure, whereas Dimension 2 demonstrated diffuse structural abnormalities and greater cognitive impairment. One *de novo* independent genetic variant was significantly associated with Dim1 but not with Dim2. Notably, the two dimensions manifested significant genetic heritability in the general population, and Dim2 was longitudinally more vulnerable to Alzheimer's disease and brain aging than Dim1.

Meanings: The two dimensions encompass heterogeneity in LLD and offer the potential for clinical precision in diagnosis and prognosis.

ABSTRACT

Importance: Late-life depression (LLD) is characterized by considerable heterogeneity in clinical manifestation. Unraveling such heterogeneity would aid in elucidating etiological mechanisms and pave the road to precision and individualized medicine.

Objective: We sought to delineate, cross-sectionally and longitudinally, disease-related heterogeneity in LLD linked to neuroanatomy, cognitive functioning, clinical symptomatology, and genetic profiles.

Design & setting: The iSTAGING study is an international multicenter consortium investigating brain aging in pooled and harmonized data from 13 studies with over 35,000 participants, including a subset of individuals with major depressive disorder.

Participants: Multimodal data from a multicentre sample (*N*=996), including neuroimaging, neurocognitive assessments, and genetics: 501 LLD participants (332 women, mean age 67.39 $\pm$ 5.56 years) and 495 healthy controls (333 women, mean age 66.53 $\pm$ 5.16 years) were analyzed. A semi-supervised clustering method (HYDRA) was applied to regional grey matter (GM) brain volumes to derive dimensional representations.

Exposure: None

Main outcome and Measure: Two dimensions were identified, which accounted for the LLD-related heterogeneity in voxel-wise GM maps, white matter (WM) fractional anisotropy (FA), neurocognitive functioning, clinical phenotype, and genetics.

Results: Dimension one (Dim1) demonstrated relatively preserved brain anatomy without WM disruptions relative to healthy controls. In contrast, dimension two (Dim2) showed widespread brain atrophy and WM integrity disruptions, along with cognitive impairment and higher depression severity. Moreover, one *de novo* independent genetic variant (rs13120336) was

significantly associated with Dim 1 but not with Dim 2. Notably, the two dimensions demonstrated significant SNP-based heritability of 18-27% within the general population (*N*=12,518 in UKBB). Lastly, in a subset of individuals having longitudinal measurements, Dim2 demonstrated a more rapid longitudinal change in GM and brain age, and was more likely to progress to Alzheimer's disease, compared to Dim1 (*N*=1,431 participants and 7,224 scans from ADNI, BLSA, and BIOCARD datasets).

Conclusions and Relevance: Heterogeneity in LLD was represented by two dimensions with distinct neuroanatomical, cognitive, clinical, and genetic profiles. This dimensional approach provides a novel mechanism for investigating the heterogeneity of LLD and the relevance of the latent dimensions to possible disease mechanisms, clinical outcomes, and responses to interventions.

Introduction

Major depressive disorder (MDD) is one of the most common mental health disorders and is a leading contributor to disability worldwide (1, 2). Late-life depression (LLD) generally refers to MDD that is present from 60-65 years of age, which can be early-onset or late-onset, affecting 1.8-7.2% of older adults in the general community (3, 4).

There is considerable heterogeneity in the clinical presentation and illness progression (5, 6). Pharmacological and psychological treatments tend to be less effective in LLD. Up to 50% of LLD patients do not achieve remission with their first treatment (7). LLD is associated with cognitive impairments (5,6) and high comorbidity, including cardiac and cerebrovascular disease (8), stroke (9), as well as increased risk for obesity, diabetes, frailty (10), and neurodegenerative diseases such as Alzheimer's disease and vascular dementia (11–14).

Magnetic resonance imaging (MRI) has revealed grey matter (GM) reductions in bilateral anterior cingulate and medial frontal cortices, insula, putamen, and globus pallidus, extending into parahippocampal gyrus, amygdala, and hippocampus. In contrast, larger GM volumes have been observed in the lingual gyrus (15), putamen and caudate regions (16). Diffusion tensor imaging (DTI) demonstrates widespread losses in white matter (WM) integrity, including in the anterior thalamic radiation, cingulum, corticospinal tract, superior and inferior longitudinal fasciculi, and uncinate fasciculus (17). Collectively, the findings support biological models of LLD being associated with cortical atrophy and white matter abnormalities in specific brain networks, although the extent and magnitude have varied.

A developing body of methodological advancement in data-driven biological subtypes (18–23) is challenging the traditional definition of neurological diseases, such as Alzheimer's disease (18, 19, 21, 22) and MDD (24–26). One of the critical advantages of semi-supervised

clustering methods (22, 23, 27, 28) is that it performs subtyping via the "*1-to-k*" mapping from the domain of a reference group (i.e., healthy control) to the patient group, thereby avoiding clustering patients according to disease-irrelevant confounds. Distinct neuropathological mechanisms may underlie heterogeneity in the presentation and progression of the clinical phenotype (29). Furthermore, the extent to which genetic heterogeneity influences or interacts with phenotypic expression has barely been explored (30), while individual-level variability, including environment, genetic or other factors, may lead to different levels of disease liability (31).

We sought to delineate heterogeneity in LLD in a large multicenter sample ($N = 996$) using a state-of-the-art semi-supervised clustering method (HYDRA)[1] (27). We hypothesized that multiple distinct dimensions coexist to account for the underlying heterogeneity and that these dimensions might be prominent in the general population and longitudinal trajectories.

Materials and methods

Participants

The iSTAGING study is an international consortium consisting of various imaging protocols, scanners, data modalities, and pathologies (32), comprising harmonized MRI data in more than 35,000 participants, from over 13 studies, and encompassing a wide range of ages (22 - 90 years). The present study includes LLD from four cohorts, including UK Biobank (UKBB) (33); Psychotherapy Response Study at the University of California San Francisco (UCSF); Baltimore Longitudinal Study of Aging (BLSA) (34, 35); and Biomarkers of Cognitive Decline Among Normal Individuals (BIOCARD). The institutional review board per site approved the study. All participants provided informed consent to the studies contributing data to this pooled meta-analysis.

[1] https://github.com/anbai106/mlni

We applied a harmonized LLD definition criterion to consolidate LLD participants and excluded comorbid medical and neurological diseases that were potential confounds (i.e., *LLD population*) as follows: i) all participants from all four sites are restricted to be 60 years or above; ii) For UKBB, we excluded subjects that are diagnosed with schizophrenia, bipolar, psychotic symptoms, anxiety, obsessive-compulsive disorder, post-traumatic stress disorder, Huntington's disease, AD, epilepsy and stroke, diabetes, or hypertension; For BLSA, we excluded individuals diagnosed with hypertension, anxiety, bipolar or schizophrenia; For BIOCARD, we excluded individuals diagnosed with diabetes or hypertension; For UCSF, we excluded individuals with substance abuse, psychotic features, cognitive-enhancing use, neurological diseases, or post-traumatic stress disorders (**Table 1**). Moreover, we defined two additional populations: i) *general population* (12,518 participants from UKBB) and ii) *longitudinal population* (1,431 participants from ADNI, BLSA, and BIOCARD). A total of 996 participants (501 LLD patients and 495 healthy control subjects) were included for the LLD population. Image protocols and acquisition parameters for all sites are presented in **Supplementary eMethod 1**.

Image preprocessing

Quality-controlled (QC) images were corrected for magnetic field intensity inhomogeneity (36) (**Supplementary eMethod 2**). A state-of-the-art multi-atlas parcellation method (MUSE) (37) was used to extract regions of interest (ROI) values of the segmented GM tissue maps (**Supplementary eTable 2**). Voxel-wise regional volumetric maps (RAVENS) for each tissue volume (38) were generated by spatially aligning the skull-stripped images to a template residing in the MNI-space using a registration method (39). Fractional anisotropy (FA) maps were used to examine microstructural integrity disruptions in WM (**Supplementary eMethod 3**)**.** The mean FA values

were extracted within the 48 WM tracts of the JHU ICBM-DTI-81 WM label atlas (40). Inter-site image harmonization of the GM MUSE ROIs is detailed in **Supplementary eMethod 4.**

Genetic preprocessing

We consolidated an imaging-genetic dataset from UKBB that passed the QC protocol, resulting in 20,438 participants and 8,430,655 single nucleotide polymorphisms (SNPs) (**Supplementary eMethod 8)**. We then selected 774 UKBB participants who overlapped with the LLD population for genetic analyses.

Discovery of the multidimensional representation via HYDRA

We applied a semi-supervised clustering method, termed HYDRA (27) (**Supplementary eMethod 5**), to the harmonized MUSE ROIs. Briefly, HYDRA aims to cluster disease effects instead of directly clustering participants by comparing the patterns between healthy controls (CN) and LDD patients, thus resulting in a "*1-to-k*" mapping from the CN to the patient domain.

We chose the optimal number of dimensions/clusters (*k*), ranging from 2 to 8 clusters, by the Adjusted Rand Index (ARI) (41). We performed additional analyses to evaluate the robustness of the optimal *k* clusters scheme. First, split-sample analyses (42) were carried out to assess whether the dimensions in each half exhibit similar neuroanatomical patterns, given that the two halves had similar cohort characteristics in terms of age, sex, and site. Secondly, we conducted leave-site-out validation (43) to examine if the dimensions were consistent across sites: i) training on UKBB only and ii) training on all sites. Lastly, a permutation test was performed to test the statistical significance with the optimal *k* cluster scheme (**Supplementary eMethod 6**).

Evaluation of the multidimensional representation in neuroimaging, cognition, and genetics

We subsequently investigated their characteristics regarding i) GM volume, ii) WM integrity, iii) cognitive functioning and depression-related variables, and iv) genetic architecture. Moreover, we investigated the expression of the *k* dimensions in the general population and longitudinal data.

Voxel-wise GM RAVENS regional tissue volumes

Voxel-wise RAVENS GM maps from all sites were used to assess the differences in GM tissue volumes. The *3dttest++* program (44) in AFNI (45) was used to detect the distinct neuroanatomical patterns of the corresponding dimensions vs. the CN group, considering age, sex, site, and ICV as covariates. Finally, for those voxels that survive the adjustment (Benjamini-Hochberg procedure), voxel-wise effect-size maps (i.e., Cohen's *f2*) were estimated for each paired comparison.

Regional WM integrity abnormality

WM microstructural abnormality was assessed using the mean FA values of the 48 regional tracts from UKBB. Group comparisons were performed with multiple linear regression models using R (version 3.4.0, The R Foundation) (**Supplementary eMethod 9**). Age and sex were fixed effects, and group was the variable of interest. P-values were corrected, and Cohen's *f2* was computed with the same procedure as above.

Demographic, cognitive, and clinical variables

Group comparisons for demographic, cognitive, and clinical variables (**Supplementary eTable 5**) were examined separately between the dimensions. Mann–Whitney–Wilcoxon test was used for continuous variables (e.g., age) and Chi-Square test of independence for categorical variables (e.g., sex). Global effect size (i.e., Cohen's *d*) was also reported for continuous variables.

214

215 ***Genome-wide associations***

216 We performed GWAS with the derived binary dimension traits, i.e., Dim1 or Dim2 vs. CN using
217 Plink 2[2]. FUMA online platform[3] was then used to annotate the genomic risk loci and independent
218 significant SNPs (**Supplementary eMethod 8**).

219

220 ***Evaluation of the multiple dimensions in the general population***

221 The trained model was applied to the external validation samples in the general population (Table
222 1). Dimension membership (**Fig. 3B**) and expression scores of the *k* dimensions were derived
223 (**Supplementary eMethod 7)**.

224 We examined the neuroanatomical patterns using RAVENS GM maps, demographic and
225 cognitive functioning of the *k* dimensions in the general population. We calculated the genome-
226 wide SNP-based heritability coefficient (h^2) using GCTA[4] (**Supplementary eMethod 8**).

227

228 ***Evaluation of the multiple dimensions in longitudinal data and their progress to AD and brain***
229 ***aging***

230 The cross-sectionally trained model was applied to the longitudinal population (**Table 1**).
231 Dimension membership was derived to evaluate its longitudinal changes in MUSE GM ROIs,
232 SPARE-AD (Spatial Patterns of Atrophy for REcognition of AD) (46), SPARE-BA (Brain Age)
233 (47). Specifically, the Rate of Change (RC) over time in these variables for each participant was
234 derived with a linear mixed-effects model and compared across dimensions using a linear
235 regression model (**Supplementary eMethods 9, 10**).

236

[2] https://www.cog-genomics.org/plink/2.0/
[3] https://fuma.ctglab.nl/
[4] https://cnsgenomics.com/software/gcta

Results

HYDRA reveals two dimensions

The highest ARI (0.58) was achieved by a HYDRA model for k=2 clusters (**Supplementary eFigure 1**). The cluster assignment distribution for k = 2 to 8 across sites is presented in **Supplementary eTable 3**. For the optimal k=2 clustering scheme, 227 LLD participants were assigned to Dimension 1 (Dim1) and 274 to Dimension 2 (Dim2). The optimal k=2 clustering scheme was replicated in split-sample and leave-site-out analyses (**Supplementary eFigure1**). In the leave-site-out analyses, the percentage overlap for participants assigned to the same dimension was 89.12% (91.77% for UKBB, 76.41% for BLSA, 81.27% for BIOCARD, and 84.45% for UCSF). The neuroanatomical patterns of the two dimensions were similar (**Supplementary eFigure 3**) to the original dimension patterns (Fig. 1). In split-sample analyses, the GM patterns for the two splits were similar (**Supplementary eFigure 2**) compared to the original dimension patterns (**Fig. 1A**). The ARI at k=2 was higher than the null distribution in the permutation test (P-value<0.001). Lastly, we presented the results without excluding comorbidities in UKBB, which yielded similar imaging patterns for the two dimensions (**Supplementary eFigure 4**). Therefore, we present the results of k=2 for all subsequent analyses.

Differences in GM volumetric patterns

Dim1 demonstrated greater GM tissue volume in bilateral thalamus, putamen, and caudate relative to healthy controls. Dim2 demonstrated reduced GM tissue volume in widespread cortical regions, including bilateral anterior and posterior cingulate gyri, superior, middle, and inferior frontal gyri, gyrus recti, insular cortices, superior, middle, and inferior temporal gyri, etc., compared to controls

(**Fig. 1A**). The split-sample and leave-site-out analyses are detailed in **Supplementary eFigure 2** and **Supplementary eFigure 3**, respectively.

Differences in WM integrity disruption

Dim1 exhibited similar FA values compared to controls. However, Dim2 showed widespread WM disruptions, with 31 out of the 48 WM tracts demonstrating significantly lower FA values than controls but small effect sizes ($0.01 \leq$ Cohen's $f2 \leq 0.05$, **Fig. 1B**). Specifically, the middle cerebellar peduncle tract obtained the highest effect size (Cohen's $f2$=0.05). Other affected WM tracts mainly involved frontal lobe and subcortical limbic regions (**Supplementary eTable 4**).

Dim1 and Dim2 demonstrate differences in clinical profiles

Dim1 showed statistically higher scores in Fluid Intelligence scores (Cohen's d = 0.25) and fewer errors in Pairs Matching test (Cohen's d = -0.28), and fewer depressive symptoms in Patient Health Questionnaire responses (PHQ9) (Cohen's d = -0.45) relative to Dim2. The two dimensions did not significantly differ in age, sex, site, or other clinical variables (details in **Supplementary eTable 5**).

Differences in genome-wide associations

Dim1, but not Dim2, was significantly associated with one *de novo* independent variant (rs13120336 on chromosome 4) (P-value=3.14×10^{-8}) (**Fig. 2**). Quantile-quantile plots are presented in **Supplementary eFigure 5**.

Expression of the two dimensions in the general population

Applying the trained model to UKBB samples resulted in: 2269 Dim1 participants, 3786 Dim2 participants, and 2963 Mixed individuals (both dimensions were expressed), and 3500 None participants (neither dimension was expressed) (**Supplementary eTable 6** and **Fig. 3B**).

The neuroanatomical patterns of the two dimensions were stable (**Fig. 3A**). Dim1 showed higher scores in Fluid Intelligence scores (P-value < 1e-10, Cohen's d = 0.28), but lower errors in Pairs Matching (P-value < 1e-6, Cohen's d = -0.13) compared to Dim2 (**Supplementary eTable 6**). The expression scores of the two dimensions were significantly heritable in the general population. Specifically, h^2 for Dim1 and Dim2 was 0.27 ± 0.04 (P-value<5.7e-10), and 0.18 ± 0.04 (P-value<1.1e-5), respectively.

The two dimensions and longitudinal trajectories

Applying the trained model to ADNI, BLSA, and BIOCARD, which also had longitudinal follow-up data, yielded 301 Dim1 participants, 390 Dim2 participants, 330 Mixed individuals, and 410 None participants in baseline images (**Supplementary eTable 7**).

The neuroanatomical patterns of the two dimensions were stable (**Fig. 4A**). The GM RC in Dim2 decreased more rapidly than Dim1 or None groups (-0.1 < Cohen's $f2$ < 0.1), specifically in the left precentral gyrus, temporal pole, and right anterior insula (**Fig. 4B**). Moreover, the two dimensions remained independent and stable along longitudinal trajectories (**Fig. 4C**). Lastly, Dim2 showed progression of both SPARE-AD (Cohen's $f2$=0.03) and SPARE-BA (Cohen's $f2$=0.03) compared to Dim1 (**Fig. 4D**), but not at baseline.

Discussion

Two reproducible and distinct dimensions characterized neuroanatomical heterogeneity in LLD. Dim1 showed relatively preserved brain anatomy with larger subcortical regional volumes and was associated with one *de novo* genetic variant, while Dim2 displayed widespread brain atrophy and WM integrity disruptions with impaired cognitive functioning and increased depressive severity. Moreover, the two dimensions were manifested in the general population and were significantly heritable. Notably, Dim2 demonstrated a higher degree of progression to AD and brain aging signatures relative to Dim1.

The two dimensions demonstrate the extent of underlying GM heterogeneity in patients with LLD. GM atrophy evident in Dim2 has been widely reported in previous case-control studies (48–50). Regional atrophy in the frontal lobes has been observed (51, 52), which is associated with cognitive deficits as well as reports of psychotic symptoms (53). Striatal atrophy has been associated with degeneration in the dopaminergic connections between caudate and cortical limbic areas involved in mood regulation (54), although increased caudate and putamen volumes have been found in UKBB depression phenotypes (16). Dim2 showed atrophy in hippocampal regions, perhaps indicative of future neuro progressive degeneration linked with Alzheimer's disease.

The two identified neuroanatomical dimensions differed significantly in microstructural integrity. Dim1 shows no significant WM abnormalities, while Dim2 demonstrates widespread WM abnormalities. WM lesions may play a key role in conferring vulnerability or perpetuating depressive syndromes in LLD and contributing to the observed microstructural disturbance (55). Widespread WM disruptions can persist in LLD, even excluding WM lesions from the DTI analysis (56). WM tracts connecting fronto-subcortical and fronto-limbic regions are most frequently affected, including the uncinate fasciculus (57, 58), anterior thalamic radiation, superior

longitudinal fasciculus (55, 57, 59), and posterior cingulate cortex (60). Dim2 demonstrates clinical features of LLD patients that are frequently associated with more severe cognitive deterioration (61–63). Interestingly, previous studies using depressive symptom and cognitive scores (25), or metabolic-inflammatory profile (26), derived one subtype that was a 'healthy' group, and other subgroups that demonstrated higher depressive symptom scores or a more specific immune-inflammatory dysregulation profile.

The detected genetic variant (rs13120336) was uniquely associated with Dim1. Interestingly, two mapped genes (CCDC110 and LOC105377590) have been previously linked to cancer and diabetes (64, 65). We speculate that these genetic factors may play a key role in the heterogeneity of imaging phenotype and cognitive dysfunctions in the two dimensions. Many studies have shown that depression is associated with different genetic variants, some of which were not replicated (66–69). Replication needs to be performed to confirm this detected variant. In general, our dimensional approach might provide another approach for genetic associations in depression.

The two dimensions showed significant genetic heritability of 18-27%, potentially suggesting genetic underpinnings of neuroanatomical phenotypes associated with depression in the general population. Of note, multimorbidity, such as schizophrenia and anxiety disorders, exists in the UKBB population (70). Such comorbidities might account for the expression of the two dimensions to some extent. MDD is a common and complex syndrome with an estimated genetic heritability of approximately 40% (71), and prevalence rates range from 7 - 13% (69). Our findings confirm the high risks and prevalence of depression in the general population.

The proposed two-dimensional representation emphasizes the tremendous prognostic potential to distinguish LLD that is co-occurring or preceding neurodegenerative diseases. Dim2

progressed towards an AD or brain aging signature, whereas Dim1 expressed a preserved brain anatomy. Epidemiological studies (72, 73) have consistently found that shared risk factors exist in AD and LLD, supporting depression as a prodromal feature or a risk factor of AD. Interestingly, the two dimensions did not longitudinally differ in cognitive impairment, perhaps supporting the AD pathological cascade model (74).

To ensure the reproducibility of the finding, we had performed additional analyses: split-sample analysis, leave-site-out analysis, and applying the model trained on LLD to independent UKBB and a combined ADNI, BLSA, and BIOCARD cohort with the same age range as the LLD population. From a technical perspective, applying the trained LLD model to a younger population would be possible, but this could lead to a trivial solution due to the significant difference in age ranges, rather than due to a disease effect of interest, as aging might play a crucial role in driving these dimensions. We believe that applying the model to external data requires careful consideration of potential confounds, such as demographic differences.

There are several limitations. We had sought to limit potential confounds in population selection in order to aid interpretation of the dimensions. However, this could then potentially limit generalizability of the findings. That the dimensions were reproduced in various conditions demonstrates the robustness of the dimensions. Nonetheless, longitudinal LLD data are required to confirm the added value of the proposed multidimensional representation and replication of the GWAS findings is necessary.

Conclusions

LLD was characterized by two dimensions linked to neuroanatomy, cognitive functioning and genetic profiles. The two-dimensional representation offers a system for future research on the

373 underlying etiology mechanisms, heterogeneity of genetic architectures, and the potential for
374 personalized clinical care.

Conflicts of interest

DAW served as Site PI for studies by Biogen, Merck, and Eli Lilly/Avid. He has received consulting fees from GE Healthcare and Neuronix. He is on the DSMB for a trial sponsored by Functional Neuromodulation. Dr. Saykin receives support from multiple NIH grants (P30 AG010133, P30 AG072976, R01 AG019771, R01 AG057739, U01 AG024904, R01 LM013463, R01 AG068193, T32 AG071444, and U01 AG068057 and U01 AG072177). He has also received support from Avid Radiopharmaceuticals, a subsidiary of Eli Lilly (in kind contribution of PET tracer precursor); Bayer Oncology (Scientific Advisory Board); Siemens Medical Solutions USA, Inc. (Dementia Advisory Board); Springer-Nature Publishing (Editorial Office Support as Editor-in-Chief, Brain Imaging and Behavior).

Acknowledgments

This work was supported, in part, by NIH grants 1RF1-AG054409-01 and U01-AG068057. ADNI is supported by NIH grants U01-AG024904 and RC2-AG036535. The BIOCARD study is supported by NIH grant U19-AG033655. The BLSA is supported by the Intramural Research Program, National Institute on Aging, and Research and Development Contract HHSN-260-2004-00012C. This research has been conducted using the UK Biobank Resource under Application Number 35148. The funder/sponsor had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Dr. Wen and Dr. Davatzikos had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis.

Authors' contributions:

Dr. Wen and Dr. Davatzikos take full responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Davatzikos, Wen, Fu

Acquisition, analysis, or interpretation of data: Davatzikos, Wen, Fu, Tosun, Resnick

Drafting of the manuscript: Wen, Fu

Critical revision of the manuscript for important intellectual content: Wen, Fu, Tosun, Veturi, Yang, Mamourian, Srinivasan, Skampardoni, Singh, Nawani, Bao, Erus, Shou, Abdulkadir, Habes, Doshi, Varol, Mackin, Sotiras, Fan, Sheline, Saykin, Shen, Ritchie, Wolk, Albert, Resnick, Davatzikos

Statistical and genetic analysis: Wen

Obtained funding: Davatzikos

Administrative, technical, or material support: Davatzikos

Study supervision: Davatzikos, Fu

Figure legend

Figure 1: **The two dimensions show distinct structural patterns**. Effect size maps were identified in Dimension 1 (Dim1) and Dimension 2 (Dim2) compared to controls (CN), respectively. **A**) Multiple selective views are shown in different views. Warmer color denotes brain atrophy (i.e., CN > Dim), and cooler color represents larger tissue volume (i.e., Dim > CN). Both directions are shown for each dimension. L: left; R: right. The effect size map is shown in a radiological fashion, i.e., the brain's left is shown to the right of the display. **B**) Dim1 and Dim2 demonstrate two distinct WM patterns based on FA values. Dim1 exhibits a normal appearance, without significant difference from controls; whereas Dim2 shows widespread disruptions in WM integrity. The P-value and effect size for all the 48 WM tracts are shown in **Supplementary eTable 4**. Both directions of the comparisons are performed, but effect sizes only show WM integrity disruptions. For references, Cohen's *f2* of ≥ 0.02, ≥ 0.15, and ≥ 0.35 signify small, moderate, and large effect sizes, respectively. Of note, we would like to clarify we do not claim that voxel-based differences provide validation of clustering. We simply show these comparisons to elucidate the characteristics of the dimensions determined by the machine learning algorithm so that we can appreciate the features which were found by the algorithm to be essential for the definition of these dimensions.

Figure 2: **Dim1 and Dim2 demonstrate distinct genetic profiles in GWAS**. A) Dim1 was significantly associated with a novel genomic risk locus. This significant independent SNP (rs13120336) is in LD with other seven-candidate SNPs that passed the GWAS P-value threshold (5e-8). FUMA identified two corresponding protein-encoding genes: CCDC110 and LOC105377590; B) Dim2 was not significantly associated with any variants.

Figure 3: **The expression of the two dimensions in the general population**. **A**) The two neuroanatomical dimensions in UKBB show distinct grey matter abnormalities. Effect size maps of GM patterns were identified in Dimension 1 (Dim1) and Dimension 2 (Dim2) compared to None (the dimension that does not express in Dim1 and Dim2), respectively. Multiple selective views are shown with the number of slices in the axial view. 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 *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 is shown to the right of the display. We include age, sex, and ICV as fixed effects and group (None vs. Dim1 or Dim2) as the variable of interest. The likelihood ratio test was used to test each effect. **B**) The quadrant plot after applying the HYDRA model trained on the LLD population to the external UKBB individuals. X-axis and Y-axis represent the expression scores for each individual at the Dim1 and Dim2, respectively. The dimension membership was decided based on the two expression scores, E1 and E2. Specifically, the individual was assigned as None when E1 and E2 are smaller than -0.3, as Dim1 when E1 > 0.3 and E2 < -0.3, as Dim2 when E1 < -0.3 and E2 > 0.3, and as Mixed for the other individuals.

Figure 4: **The two dimensions and longitudinal trajectories to aging and AD**. **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. **C**) Applying the HYDRA model to all available longitudinal scans with at least 6 years follow-ups. The two dimensions stay stable over time and are independent of each other. **D**) The positive RC for SPARE-AD and SPARE-BA of Dim2 is bigger than Dim1, meaning that Dim2 is more vulnerable to AD and brain aging longitudinally. Only subjects that have at least 6 time points were included for this analysis.

Table 1. Study cohort characteristics.

	LLD population			General population	Longitudinal population
	CN	LLD	P-value	CN[1]	CN[2]
N	495	501		12518	1431
Age (year) [min/max]	66.26 [60, 91.47]	67.33 [60, 91]	0.34	67.23 [60, 80]	71.88 [60, 93]
Sex/ female, %	333/ 67%	332/ 66%	0.78	6123/49%	666/47%
Education (year)	14.76± 2.68	14.87± 2.62	0.55	16.90± 2.81	16.86± 2.57
Systole	135.03± 16.83	134.75± 16.56	0.52	140.97± 18.88	124.06± 2.57
Diastole	75.59± 9.24	79.05± 9.15	0.45	82.26± 10.45	69.93± 11.05
Age of onset (year)	NA	34.62± 15.70	NA	NA	NA

Age is shown with mean and its range. 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; NA: not applicable; P: P-value. More details of the LLD population per site are presented in Supplementary eMethod 1. For the general population, we included all individuals from UKBB over 60 years old (excluded overlapping individuals in the LLD population). [1]Note that this population is cognitively healthy (CN) but might be diagnosed with other general disorders historically (ICD-10: https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=41202). More details of the general population are presented in Supplementary eTable 6. For the longitudinal population, we included all healthy controls from ADNI, BLSA, and BIOCARD that were diagnosed as CN at baseline. We present here only baseline information[2]. For more details, refer to Supplementary eTable 7.

References

1. Belmaker RH, Agam G: Major Depressive Disorder. New England Journal of Medicine 2008; 358:55–68

2. Thornicroft G, Chatterji S, Evans-Lacko S, et al.: Undertreatment of people with major depressive disorder in 21 countries. The British Journal of Psychiatry 2017; 210:119–124

3. Beekman AT, Copeland JR, Prince MJ: Review of community prevalence of depression in later life. Br J Psychiatry 1999; 174:307–311

4. Luppa M, Sikorski C, Luck T, et al.: Age- and gender-specific prevalence of depression in latest-life--systematic review and meta-analysis. J Affect Disord 2012; 136:212–221

5. Alexopoulos GS: Mechanisms and treatment of late-life depression. Transl Psychiatry 2019; 9:188

6. Brodaty H, Luscombe G, Peisah C, et al.: A 25-year longitudinal, comparison study of the outcome of depression. Psychol Med 2001; 31:1347–1359

7. Gutsmiedl K, Krause M, Bighelli I, et al.: How well do elderly patients with major depressive disorder respond to antidepressants: a systematic review and single-group meta-analysis. BMC Psychiatry 2020; 20:102

8. Daskalopoulou M, George J, Walters K, et al.: Depression as a Risk Factor for the Initial Presentation of Twelve Cardiac, Cerebrovascular, and Peripheral Arterial Diseases: Data Linkage Study of 1.9 Million Women and Men. PLoS One 2016; 11:e0153838

9. Pan A, Sun Q, Okereke OI, et al.: Depression and risk of stroke morbidity and mortality: a meta-analysis and systematic review. JAMA 2011; 306:1241–1249

10. Buigues C, Padilla-Sánchez C, Garrido JF, et al.: The relationship between depression and frailty syndrome: a systematic review. Aging Ment Health 2015; 19:762–772

11. Adler G, Chwalek K, Jajcevic A: Six-month course of mild cognitive impairment and affective symptoms in late-life depression. Eur Psychiatry 2004; 19:502–505

12. Diniz BS, Butters MA, Albert SM, et al.: Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. Br J Psychiatry 2013; 202:329–335

13. Galts CPC, Bettio LEB, Jewett DC, et al.: Depression in neurodegenerative diseases: Common mechanisms and current treatment options. Neuroscience & Biobehavioral Reviews 2019; 102:56–84

14. Reynolds CF, Lenze E, Mulsant BH: Chapter 23 - Assessment and treatment of major depression in older adults [Internet], inDekosky ST, Asthana S, editorsHandbook of

514 Clinical Neurology. Elsevier, 2019, pp 429–435.[cited 2021 Apr 21] Available from:
515 https://www.sciencedirect.com/science/article/pii/B9780128047668000236

516 15. Du M: Brain grey matter volume alterations in late-life depression – Journal of Psychiatry
517 & Neuroscience [Internet]2014; [cited 2021 Apr 23] Available from: http://jpn.ca/vol39-
518 issue6/39-6-397/

519 16. Harris MA, Cox SR, Nooij L de, et al.: The Influence of Phenotyping Method on
520 Structural Neuroimaging Associations with Depression in UK Biobank [Internet].
521 2020[cited 2021 Sep 11] Available from:
522 https://www.medrxiv.org/content/10.1101/2020.12.18.20248488v1

523 17. Wen M-C, Steffens DC, Chen M-K, et al.: Diffusion tensor imaging studies in late-life
524 depression: systematic review and meta-analysis. International Journal of Geriatric
525 Psychiatry 2014; 29:1173–1184

526 18. Young AL, The Alzheimer's Disease Neuroimaging Initiative (ADNI), Young AL, et al.:
527 Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases
528 with Subtype and Stage Inference. Nat Commun 2018; 9:4273

529 19. Zhang X, Mormino EC, Sun N, et al.: Bayesian model reveals latent atrophy factors with
530 dissociable cognitive trajectories in Alzheimer's disease. Proc Natl Acad Sci USA 2016;
531 113:E6535–E6544

532 20. Eshaghi A, Young AL, Wijeratne PA, et al.: Identifying multiple sclerosis subtypes using
533 unsupervised machine learning and MRI data. Nat Commun 2021; 12:2078

534 21. Vogel JW, Young AL, Oxtoby NP, et al.: Four distinct trajectories of tau deposition
535 identified in Alzheimer's disease. Nat Med 2021; 27:871–881

536 22. Yang Z, Nasrallah IM, Shou H, et al.: Disentangling brain heterogeneity via semi-
537 supervised deep-learning and MRI: dimensional representations of Alzheimer's Disease
538 [Internet]. arXiv:210212582 [cs, eess, q-bio] 2021; [cited 2021 Apr 9] Available from:
539 http://arxiv.org/abs/2102.12582

540 23. Wen J, Varol E, Sotiras A, et al.: Multi-scale semi-supervised clustering of brain images:
541 Deriving disease subtypes. Med Image Anal 2021; 75:102304

542 24. Drysdale AT, Grosenick L, Downar J, et al.: Resting-state connectivity biomarkers define
543 neurophysiological subtypes of depression. Nature Medicine 2017; 23:28–38

544 25. Lugtenburg A, Zuidersma M, Wardenaar KJ, et al.: Subtypes of Late-Life Depression: A
545 Data-Driven Approach on Cognitive Domains and Physical Frailty. J Gerontol A Biol Sci
546 Med Sci 2021; 76:141–150

547 26. Kokkeler KJE, Marijnissen RM, Wardenaar KJ, et al.: Subtyping late-life depression
548 according to inflammatory and metabolic dysregulation: a prospective study.
549 Psychological Medicine 2020; 1–11

27. Varol E, Sotiras A, Davatzikos C: HYDRA: Revealing heterogeneity of imaging and genetic patterns through a multiple max-margin discriminative analysis framework. NeuroImage 2017; 145:346–364

28. Dong A, Honnorat N, Gaonkar B, et al.: CHIMERA: Clustering of Heterogeneous Disease Effects via Distribution Matching of Imaging Patterns. IEEE Trans Med Imaging 2016; 35:612–621

29. Fu CHY, Fan Y, Davatzikos C: Addressing heterogeneity (and homogeneity) in treatment mechanisms in depression and the potential to develop diagnostic and predictive biomarkers. Neuroimage Clin 2019; 24:101997

30. Tsang RSM, Mather KA, Sachdev PS, et al.: Systematic review and meta-analysis of genetic studies of late-life depression. Neurosci Biobehav Rev 2017; 75:129–139

31. Cai N, Choi KW, Fried EI: Reviewing the genetics of heterogeneity in depression: operationalizations, manifestations and etiologies. Human Molecular Genetics 2020; 29:R10–R18

32. Habes M, Pomponio R, Shou H, et al.: The Brain Chart of Aging: Machine-learning analytics reveals links between brain aging, white matter disease, amyloid burden, and cognition in the iSTAGING consortium of 10,216 harmonized MR scans. Alzheimer's & Dementia 2021; 17:89–102

33. 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

34. 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

35. 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

36. Tustison NJ, Avants BB, Cook PA, et al.: N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 2010; 29:1310–1320

37. Doshi J, Erus G, Ou Y, et al.: MUSE: MUlti-atlas region Segmentation utilizing Ensembles of registration algorithms and parameters, and locally optimal atlas selection. Neuroimage 2016; 127:186–195

38. Davatzikos C, Genc A, Xu D, et al.: Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. Neuroimage 2001; 14:1361–1369

39. Ou Y, Sotiras A, Paragios N, et al.: DRAMMS: Deformable Registration via Attribute Matching and Mutual-Saliency Weighting. Med Image Anal 2011; 15:622–639

40. Hua K, Zhang J, Wakana S, et al.: Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. Neuroimage 2008; 39:336–347

41. Hubert L, Arabie P: Comparing partitions. Journal of Classification 1985; 2:193–218

42. Ben-Hur A, Elisseeff A, Guyon I: A stability based method for discovering structure in clustered data. Pac Symp Biocomput 2002; 6–17

43. Arlot S, Celisse A: A survey of cross-validation procedures for model selection. Statistics Surveys 2010; 4:40–79

44. Cox RW, Chen G, Glen DR, et al.: fMRI clustering and false-positive rates. Proc Natl Acad Sci U S A 2017; 114:E3370–E3371

45. Cox RW: AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 1996; 29:162–173

46. Davatzikos C, Xu F, An Y, et al.: Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. Brain 2009; 132:2026–2035

47. Habes M, Erus G, Toledo JB, et al.: White matter hyperintensities and imaging patterns of brain ageing in the general population. Brain 2016; 139:1164–1179

48. Andreescu C, Butters MA, Begley A, et al.: Gray matter changes in late life depression--a structural MRI analysis. Neuropsychopharmacology 2008; 33:2566–2572

49. Ballmaier M, Narr KL, Toga AW, et al.: Hippocampal morphology and distinguishing late-onset from early-onset elderly depression. Am J Psychiatry 2008; 165:229–237

50. Bell-McGinty S, Butters MA, Meltzer CC, et al.: Brain morphometric abnormalities in geriatric depression: long-term neurobiological effects of illness duration. Am J Psychiatry 2002; 159:1424–1427

51. Lavretsky H, Ballmaier M, Pham D, et al.: Neuroanatomical Characteristics of Geriatric Apathy and Depression: A Magnetic Resonance Imaging Study. Am J Geriatr Psychiatry 2007; 15:386–394

52. Lavretsky H, Roybal DJ, Ballmaier M, et al.: Antidepressant exposure may protect against decrement in frontal gray matter volumes in geriatric depression. J Clin Psychiatry 2005; 66:964–967

53. Elderkin-Thompson V, Hellemann G, Pham D, et al.: Prefrontal brain morphology and executive function in healthy and depressed elderly. Int J Geriatr Psychiatry 2009; 24:459–468

54. Krishnan KR, McDonald WM, Escalona PR, et al.: Magnetic resonance imaging of the caudate nuclei in depression. Preliminary observations. Arch Gen Psychiatry 1992; 49:553–557

55. Dalby RB, Frandsen J, Chakravarty MM, et al.: Depression severity is correlated to the integrity of white matter fiber tracts in late-onset major depression. Psychiatry Res 2010; 184:38–48

56. Shimony JS, Sheline YI, D'Angelo G, et al.: DIFFUSE MICROSTRUCTURAL ABNORMALITIES OF NORMAL APPEARING WHITE MATTER IN LATE LIFE DEPRESSION: A DIFFUSION TENSOR IMAGING STUDY. Biol Psychiatry 2009; 66:245–252

57. Sexton CE, Allan CL, Le Masurier M, et al.: Magnetic Resonance Imaging in Late-Life Depression: Multimodal Examination of Network Disruption [Internet]. Arch Gen Psychiatry 2012; 69[cited 2021 Mar 28] Available from: http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/archgenpsychiatry.2011.1862

58. Taylor WD, MacFall JR, Gerig G, et al.: Structural integrity of the uncinate fasciculus in geriatric depression: Relationship with age of onset. Neuropsychiatr Dis Treat 2007; 3:669–674

59. Shen X, Reus LM, Cox SR, et al.: Subcortical volume and white matter integrity abnormalities in major depressive disorder: findings from UK Biobank imaging data. Sci Rep 2017; 7:5547

60. Alves GS, Karakaya T, Fußer F, et al.: Association of microstructural white matter abnormalities with cognitive dysfunction in geriatric patients with major depression. Psychiatry Res 2012; 203:194–200

61. Rhodes E, Insel PS, Butters MA, et al.: The Impact of Amyloid Burden and APOE on Rates of Cognitive Impairment in Late Life Depression. J Alzheimers Dis 2021;

62. Wilkins CH, Mathews J, Sheline YI: Late life depression with cognitive impairment: Evaluation and treatment. Clin Interv Aging 2009; 4:51–57

63. de Nooij L, Harris MA, Adams MJ, et al.: Cognitive functioning and lifetime major depressive disorder in UK Biobank. Eur Psychiatry 2020; 63:e28

64. Lee SN, Hong K-M, Seong YS, et al.: Ectopic Overexpression of Coiled-Coil Domain Containing 110 Delays G2/M Entry in U2-OS Cells. Dev Reprod 2020; 24:101–111

65. Monji M, Nakatsura T, Senju S, et al.: Identification of a novel human cancer/testis antigen, KM-HN-1, recognized by cellular and humoral immune responses. Clin Cancer Res 2004; 10:6047–6057

66. Howard DM, Adams MJ, Clarke T-K, et al.: Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions2019; 33

67. Wray NR, Ripke S, Mattheisen M, et al.: Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nature Genetics 2018; 50:668–681

68. Shen X, Howard DM, Adams MJ, et al.: A phenome-wide association and Mendelian Randomisation study of polygenic risk for depression in UK Biobank. Nature Communications 2020; 11:2301

69. Lim GY, Tam WW, Lu Y, et al.: Prevalence of Depression in the Community from 30 Countries between 1994 and 2014. Sci Rep 2018; 8:2861

70. Dönertaş HM, Fabian DK, Fuentealba M, et al.: Common genetic associations between age-related diseases. Nat Aging 2021; 1:400–412

71. Corfield EC, Yang Y, Martin NG, et al.: A continuum of genetic liability for minor and major depression. Transl Psychiatry 2017; 7:e1131–e1131

72. Kida J, Nemoto K, Ikejima C, et al.: Impact of Depressive Symptoms on Conversion from Mild Cognitive Impairment Subtypes to Alzheimer's Disease: A Community-Based Longitudinal Study. J Alzheimers Dis 2016; 51:405–415

73. Ownby RL, Crocco E, Acevedo A, et al.: Depression and risk for Alzheimer disease: systematic review, meta-analysis, and metaregression analysis. Arch Gen Psychiatry 2006; 63:530–538

74. Jack CR, Knopman DS, Jagust WJ, et al.: Update on hypothetical model of Alzheimer's disease biomarkers. Lancet Neurol 2013; 12:207–216