Received: 29 April 2024 | Revised: 19 November 2024 | Accepted: 25 November 2024

DOI: 10.1002/alz.14490

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Compositional brain scores capture Alzheimer's disease–specific structural brain patterns along the disease continuum

Patricia Genius[1,2,3,4] | **M. Luz Calle**[5] | **Blanca Rodríguez-Fernández**[1,2,3] | **Carolina Minguillon**[1,2,6] | **Raffaele Cacciaglia**[1,2,6] | **Diego Garrido-Martin**[7] | **Manel Esteller**[8,9,10,11] | **Arcadi Navarro**[1,3,9,12,13] | **Juan Domingo Gispert**[1,2,14,15] | **Natalia Vilor-Tejedor**[1,2,3,16] | **for the Alzheimer's Disease Neuroimaging Initiative** | **for the ALFA study**

[1]BarcelonaBeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain

[2]Hospital del Mar Research Institute, Barcelona, Spain., Barcelona, Spain

[3]Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

[4]Doctoral School, PhD programme in Bioinformatics, University of Vic–Central University of Catalonia (UVic-UCC), Vic, Spain

[5]Biosciences Department, Faculty of Sciences, Technology and Engineering, University of Vic–Central University of Catalonia, Vic, Spain

[6]Centro de Investigación Biomédica en Red de Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Madrid, Spain

[7]Department of Genetics, Microbiology and Statistics, University of Barcelona (UB), Barcelona, Spain

[8]Josep Carreras Leukaemia Research Institute (IJC), Ctra de Can Ruti, Camí de les Escoles, Barcelona, Spain

[9]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

[10]Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona (UB), Barcelona, Spain

[11]Centro de Investigación Biomédica en Red Cancer (CIBERONC), Madrid, Spain

[12]Pompeu Fabra University, Barcelona, Spain

[13]Institute of Evolutionary Biology (CSIC-UPF), Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain

[14]Centro de Investigación Biomédica en Red Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, Spain

[15]Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain

[16]Department of Human Genetics, Radboud University Medical Center, Nijmegen, the Netherlands

Correspondence
Natalia Vilor-Tejedor, BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation, C. Wellington, 30, Barcelona 08005, Spain.
Email: nvilor@barcelonabeta.org

The complete list of collaborators of the ALFA study can be found in the Acknowledgments section.

Abstract

INTRODUCTION: Traditional multivariate methods for neuroimaging studies overlook the interdependent relationship between brain features. This study addresses this gap by analyzing relative brain volumetric patterns to capture how Alzheimer's disease (AD) and genetics influence brain structure along the disease continuum.

METHODS: This study analyzed data from participants across the AD continuum from the Alzheimer's and Families (ALFA) and Alzheimer's Disease Neuroimaging Initiative (ADNI) studies. Compositional data analysis (CoDA) was exploited to examine relative

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Funding information
Alzheimer's Disease Neuroimaging Initiative; National Institutes of Health, Grant/Award Number: U01 AG024904; 'la Caixa' Foundation, Grant/Award Number: 100010434; Health Research and Innovation Strategic Plan, Grant/Award Number: SLT002/16/00201; Universities and Research Secretariat; Ministry of Business and Knowledge of the Catalan Government, Grant/Award Number: 2021 SGR 00913; Spanish Ministry of Science and Innovation—State Research Agency, Grant/Award Number: IJC2020-043216-I/MCIN/AEI/10.13039/501100011033; Spanish Research Agency, Grant/Award Numbers: MICIU/AEI/10.13039/501100011033, RYC2022-038136-I, PID2022-143106OA-I00; Alzheimer's Disease Data Initiative: William H. Gates Sr. Fellowship

brain volumetric variations that (1) were linked to different AD stages compared to cognitively unimpaired amyloid-β–negative (CU A–) individuals and (2) varied by AD genetic risk.

RESULTS: Disease stage–specific compositional brain scores were identified, differentiating CU A– individuals from those in more advanced stages. Genetic risk–stratified models revealed a broader genetic landscape affecting brain morphology in AD, beyond the well-known apolipoprotein E ε4 allele.

DISCUSSION: CoDA emerges as an alternative multivariate framework to deepen understanding of AD-related structural changes and support targeted interventions for those at higher genetic risk.

KEYWORDS
Alzheimer's disease genetic predisposition, brain imaging genetics, compositional brain score, compositional data analysis, multi phenotype analysis, neurodegeneration, polygenic risk scoring

Highlights

- Compositional data analysis (CoDA) revealed the relative variation of brain region volumes, captured in compositional brain scores, capable of discerning between cognitively unimpaired amyloid-β–negative individuals and subjects within other disease-stage groups along the Alzheimer's disease (AD) continuum.
- CoDA also uncovered the genetic vulnerability of specific brain regions at each stage of the disease along the continuum.
- CoDA is capable of integrating magnetic resonance imaging data from two different cohorts without stringent requirements for harmonization. This translates as an advantage, compared to traditional methods, and strengthens the reliability of cross-study comparisons by standardizing the data despite different labeling agreements, facilitating collaborative and large-scale research.
- The algorithm is sensitive to AD-specific effects, as the main compositional brain scores display little overlap with the age-specific compositional brain score.
- CoDA provides a more accurate analysis of brain imaging data addressing its compositional nature, which can influence the development of targeted approaches, opening new avenues for enhancing brain health.

1 | INTRODUCTION

Alzheimer's disease (AD) is a complex, multifactorial, and heterogeneous neurodegenerative disorder. In addition to accumulation of soluble amyloid beta (Aβ) in neuritic plaques, and the deposition of abnormally phosphorylated tau protein into neurofibrillary tangles,[1] AD is characterized by changes in brain structure[2,3] that follow a limbic-to-frontal pattern[4] throughout the disease continuum. It has been widely described that structural heterogeneity inter- and also intra-stage[5,6] can be shaped by the effect of known genetic[7,8] and environmental factors.[9,10] However, much remains to be understood about how these factors influence AD intermediate phenotypes along the disease continuum. Addressing these gaps requires innovative modeling approaches that can open new avenues for assessing the impact of risk factors on brain features.

Traditional methods for analyzing neuroimaging data often rely on modeling methods that assess individual brain features separately.[11–15] For instance, machine learning (ML) and deep learning (DL) frameworks have recently emerged as robust tools for identifying AD-related structural patterns with high predictive accuracy, offering significant potential in diagnostics and prognostics.[16,17] Although effective for prediction, these approaches may inadequately capture the coordinated structural changes within the brain, potentially overlooking the compositional nature of brain architecture. In neurodegenerative diseases like AD, structural changes rarely occur individually; instead, alterations in one region can influence, and are influenced by, changes in other regions.

Other techniques, such as structural covariance methods, have been developed to address these interdependencies, as they commonly examine structural covariance using independent component

analysis to detect latent sources of variability across the brain, providing insight into co-varying structures across subjects.[14,18] However, structural covariance relies on pairwise covariances between brain regions, overlooking the broader compositional and bounded nature of the data, where brain regions physically interact. This approach may miss essential information about the relative contributions of interdependent brain regions as components of a larger system, referred to as composition.

Mathematically, a composition is defined as a set of variables (components) with values in the positive orthant of the real dimension, that represent proportions of a whole and are constrained to a constant.[19] Thus, brain structures, irrespective of their segmentation level,[20,21] function as interdependent components within a compositional system restricted by the total brain volume. Compositional data analysis (CoDA) provides a complementary holistic framework that captures these interdependent components of a unified whole by assessing relative variations among multiple brain regions instead of analyzing raw volumes. CoDA examines the relative proportions between brain regions, ensuring that any changes in one region are understood in relation to changes in others, preserving the data's compositional structure, ensuring permutation, scale invariance, and sub-compositional coherence,[19,22] properties that structural covariance do not address. While CoDA is primarily applied in microbiome research,[22] few neuroimaging studies have used compositional approaches seeking more sophisticated approaches to evaluate the brain's complex structure focusing on the relative variation between brain region volumes.[23,24]

In this study, we aimed to identify relative volumetric variations in cortical and subcortical brain areas that (1) were associated with higher odds of belonging to advanced AD stages, compared to cognitively unimpaired Aβ-negative (CU A–) individuals, and (2) were influenced by genetic risk for AD within each disease stage. Using CoDA, we focused on the interconnected relationships among brain regions, capturing relative changes in compositional brain scores that reflect these interdependencies. Moreover, to distinguish AD-specific changes from those associated with normal aging, we conducted parallel analyses in CU A– individuals to capture the effect of higher chronological age in brain structure in the youngest non-pathological group, expected to display less age-related AD co-occurring competing risk factors[25] that could act as confounding variables for modeling the relationship between brain features and age. This comparison allowed us to identify which relative brain volumetric variations were distinct to AD progression, providing further specificity to our findings.

RESEARCH IN CONTEXT

1. **Systematic review**: Recent research emphasizes the importance of compositional approaches in brain imaging studies, particularly in Alzheimer's disease (AD), in which changes in one region can influence others. This study aims to identify relative brain volumetric variations in cortical and subcortical regions that (1) are linked to different AD stages compared to cognitively unimpaired amyloid-β–negative (CU A–) individuals and (2) vary by AD genetic risk.
2. **Interpretation**: Disease stage–specific compositional brain scores were identified, differentiating CU A– individuals from those in more advanced stages. These patterns showed minimal overlap with those associated with normal aging in CU A– individuals, underscoring the distinct impact of AD pathology. When genetic risk was considered, individuals with higher AD genetic risk exhibited differences in compositional brain scores compared to those with lower genetic risk, suggesting that AD pathology and genetic predisposition contribute unique patterns of brain modulation.
3. **Future directions**: Investigating the impact of gene risk factors' interaction on brain structure as well as the combined action of genetic factors and brain structure over time through longitudinal studies comprising diverse populations will significantly enhance our understanding of these dynamics.

2 | METHODS

2.1 | Study population

This study included 338 CU middle-aged (45–65 years old) participants from the ALFA+ cohort, a nested cohort of the Alzheimer's and Families study,[26] and 330 cognitively impaired (CI) individuals from the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort, a project funded by the National Institutes of Health (NIH) and launched in 2004 (adni.loni.usc.edu).[27] Individuals had available information on cortical and subcortical brain region volumes, cerebrospinal fluid (CSF) biomarkers, and genetic data. Participants were classified into A/T groups defined by their CSF biomarker profile according to the A/T framework described in Jack et al.[28] Aβ pathology positivity (A+) and tau pathology positivity (T+) were defined based on validated cut-offs.[29] Further details on the CSF sampling can be found in Methods SM1.1 in supporting information. The CSF profile of A+T+ individuals in ALFA was determined by the CSF Aβ42/40 ratio < 0.071 pg/mL and levels of CSF phosphorylated tau (p-tau)181 > 0.24 pg/mL, while in ADNI A+ was defined by CSF Aβ42 levels ≤ 1100 pg/mL. All participants from ADNI were CI, as defined by memory complaints and a Clinical Dementia Rating (CDR) ≥ 0.05. Contrary to AD patients, mild cognitive impairment (MCI) did not present complaints regarding general cognition and functional performance, and did not qualify for the diagnosis of dementia (Petersen et al.[27]). Participants were classified into four different groups based on both their clinical diagnosis and their CSF amyloid profile. Finally, the sample was covering the whole disease continuum: CU A– individuals ALFA ($N = 220$, 41 high genetic risk AD), CU Aβ–positive individuals (CU A+) ALFA ($N = 118$, 46 high genetic risk AD), MCI Aβ–positive individuals (MCI A+) ADNI ($N = 230$,

87 high genetic risk AD), AD Aβ–positive individuals (AD A+) ADNI ($N = 100$, 42 high genetic risk AD).

2.2 | Genetic data acquisition, quality control, and imputation

For the ALFA study, DNA was obtained from blood samples through a salting-out protocol. Genotyping was performed with the Illumina Infinium Neuro Consortium (NeuroChip) Array (build GRCh37/hg19).[30] A quality control procedure was performed using PLINK software. Imputation was performed using the Michigan Imputation Server with the haplotype Reference Consortium Panel (HRC r1.1 2016)[31] following default parameters and established guidelines. A full description of the genotyping, quality control, and imputation procedures is available elsewhere.[32] ADNI participants underwent genotyping using both the Human 610-Quad BeadChip and the Illumina HumanOmniExpress BeadChip (Illumina, Inc.). Quality control measures and imputation procedures closely followed the protocols established in the ALFA study. Further details are available elsewhere.[33]

2.3 | Genetic predisposition to AD: polygenic risk score computation

The polygenic risk score of AD (PRS_{AD}) was calculated using the PRSice version 2 tool.[34] Summary statistics from a recent genome-wide association study (GWAS) for AD[35] were obtained to compute the PRS_{AD}[32] (Table S1 in supporting information). The algorithm retained the single nucleotide polymorphisms (SNPs) with the smallest p value in each 250 kb window and removed SNPs that were in linkage disequilibrium ($r^2 > 0.1$). The threshold of SNP inclusion was defined at p value $< 5{\cdot}10^{-6}$. The PRS_{AD} was computed by adding up the alleles carried by participants, weighted by the SNP allele effect size from the GWAS and normalizing by the total number of alleles. The same procedure was performed to estimate the PRS of AD when excluding the apolipoprotein E (*APOE*) region ($PRS_{ADnoAPOE}$; chr19:45,409,011-45,412,650; GRCh37/hg19). PRSs were computed in ALFA and ADNI.[32] Both PRSs were dichotomized and we created two groups of subjects to differentiate individuals at higher genetic predisposition to AD from the rest of the sample. Categorization was done in each cohort separately. The cut-off point was defined by the quantile 0.8 (high genetic group: PRS ≥ quantile 0.8; low genetic group: PRS < quantile 0.8).

2.4 | Brain region segmentation and quantification

Volumes for cortical and subcortical brain regions were obtained through high-resolution 3D T1-weighted magnetic resonance imaging (MRI) scans. Each cohort used customized protocols specific to the scanner. The images were pre-processed and segmented using FreeSurfer. In ALFA,[36] protocol parameters were identical for all participants and included high-resolution 3D structural images weighted in T1 with an isotropic voxel of $0.75 \times 075 \times 0.75$ mm^3, acquired with a 3T Philips ingenia CS scanner. The acquisition parameters were repetition time (TR)/echo time (TE)/inversion time (TI) = 9.9/4.6/900 ms, flip angle = 8°, and a matrix size of $320 \times 320 \times 240$. In ADNI, MRI acquisition parameters included a TR of 2400 ms, TI of 1000 ms, flip angle of 8 degrees, and a field of view (FOV) of 24 cm. The acquisition matrix was set to $256 \times 256 \times 170$ in the x, y, and z dimensions, resulting in a voxel size of $1.25 \times 1.26 \times 1.2$ mm.[37] The MRI data underwent preprocessing steps, including reorienting images to the anterior and posterior commissure line, skull stripping, cerebellum removal, and intensity inhomogeneity correction. Finally, the data were downsampled to a matrix size of $128 \times 128 \times 128$. FreeSurfer version 7 was used for cortical and subcortical quantification in ALFA, while in ADNI, version 5 and version 6 were used in different batches, although for all MCI A+ and AD A+ included in this study, images were preprocessed with FreeSurfer version 5. Due to the scarce use of compositional methods in imaging genetics studies, there are no standardized and validated protocols for harmonizing compositional data and removing batch effects. Therefore, we proposed a three-step validation protocol (Methods SM1.2 in supporting information) and applied the neuroCombat harmonization pipeline for cross-sectional designs to adjust for potential batch effects.[38,39] Sensitivity analyses were performed to compare the compositional brain scores obtained in both scenarios, pre- and post-harmonization (Results SR1.1 in supporting information). Cortical surface parcellation was done using the Desikan–Killiany atlas.[40] Subcortical measurements were obtained working with the automatic subcortical segmentation of FreeSurfer. Volumes were globally quantified by summing the measurements of both hemispheres. A total of 41 volumes were selected combining 34 measurements from cortical regions and 7 from subcortical ones (Table S2 in supporting information).

2.5 | Compositional nature of structural brain features

Compositional data consist of a vector of positive measurements whose values are restricted by their total sum[19] (Methods SM1.3 in supporting information). Given this definition, structural brain imaging features can be understood as components of a composition, restricted by the total volume. The compositional dataset of this study consists of 41 cortical and subcortical brain region volumes, all positive measurements, that are restricted by their total sum. We applied the coda4microbiome algorithm,[22] a CoDA method initially designed for microbiome studies, used here to obtain structural brain biomarkers. The method is designed to identify a model containing the minimum number of features with the maximum predictive power of the outcome of interest. The algorithm relies on the analysis of log ratios and variable selection is addressed through cross-validation penalized regression on the "all-pairs log-ratio model," the model containing all possible pairwise log-ratios $[\frac{n!}{(k!(n-k)!)}]$. The inferred model is expressed as a linear combination of the (log-transformed) brain volumes of the selected brain regions, with the peculiarity that the sum of the coeffi-

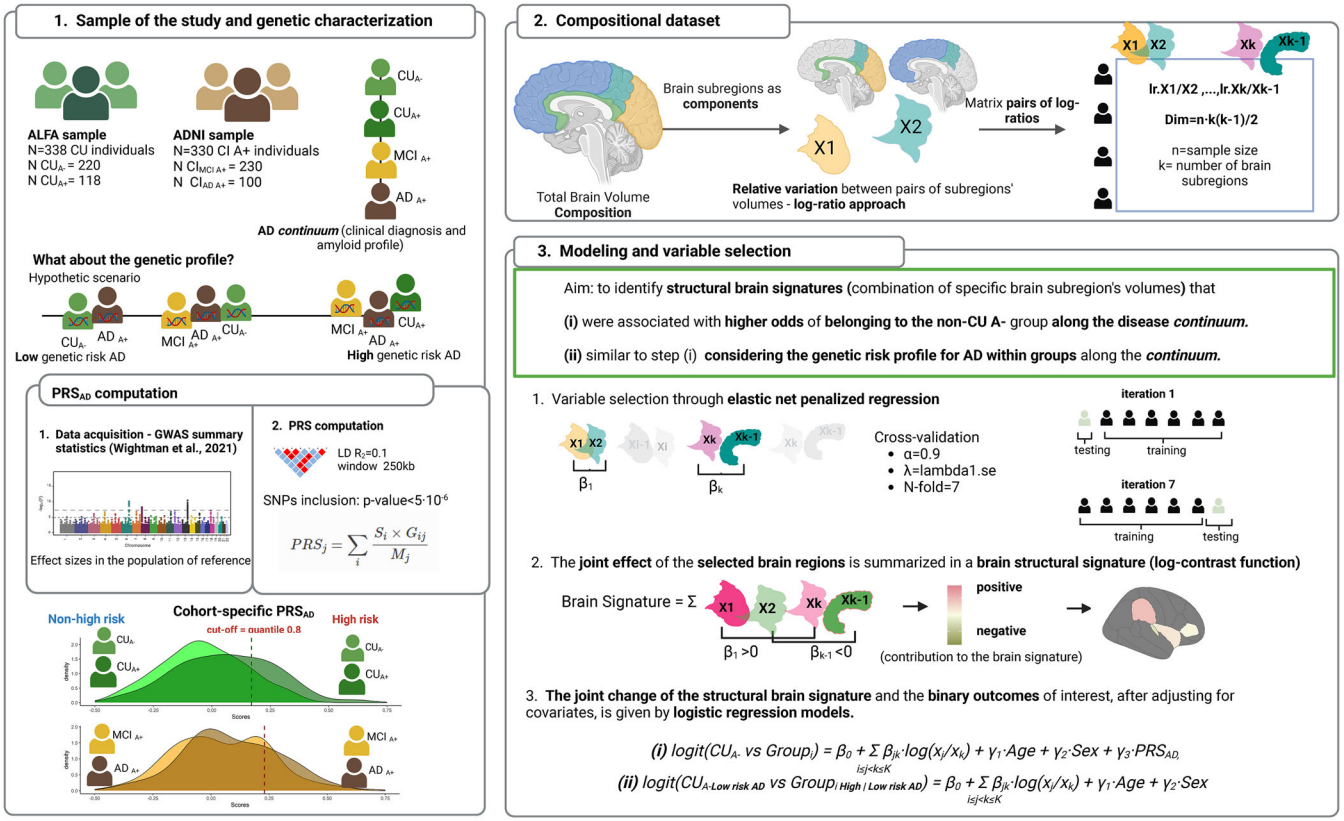

FIGURE 1 Workflow coda4microbiome algorithm: implementation in the proposed brain imaging genetics study. AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; ALFA, Alzheimer's and Families study; MCI, mild cognitive impairment; PRS, polygenic risk score; SNP, single nucleotide polymorphism.

cients is equal to zero. This defines a balance between those regions with a positive coefficient, contributing positively to the outcome, and those with a negative coefficient, that contribute negatively to the outcome.

2.6 | Statistical analyses

Differences in demographic characteristics were assessed according to amyloid status (ALFA) and diagnosis group (ADNI), using chi-square tests for categorical variables and parametric (*t* test, analysis of variance) and non-parametric tests (Wilcoxon rank sum test or Kruskal–Wallis test) for continuous normally and non-normally distributed variables, respectively. *p* values for pairwise comparisons were also provided, adjusted for Benjamini–Hochberg false discovery rate (FDR). CoDA[22,41] was performed to identify compositional brain scores (i.e., a combination of specific brain region volumes) (1) associated with higher odds of belonging to the non-CU A– group along the disease continuum, and (2) varying by genetic risk profile for AD along the disease continuum. In the first scenario, logistic regression models, adjusted for age, sex, and the PRS_{AD}, were assessed to examine the compositional brain score associated with higher odds of non-belonging to the CU A– group across the disease continuum (CU A– vs. CU A+/MCI A+/AD A+). In the second step, groups were defined based on both clinical status (CU A+, MCI A+, and AD A+) and genetic risk profile (low vs. high genetic risk for AD). All comparisons used the low genetic risk CU A– group as the reference (Figure 1). Separate analyses were performed based on the PRS calculation used to determine genetic groups: (1) the PRS including all SNPs associated with AD (PRS_{AD}) and (2) the PRS excluding the *APOE* region ($PRS_{ADnoAPOE}$). The procedure was identical for both cases, and analyses were repeated accordingly. Further, we investigated age-specific effects on brain structure by performing CoDA to explore the compositional brain score associated with higher chronological age in CU A– individuals, adjusting for sex and genetic predisposition to AD. These results were then compared to the compositional brain score associated with higher odds of belonging to the non-CU A– group along the disease continuum. Finally, we conducted post hoc analyses to assess the robustness of the variable selection procedure for generating brain compositional scores. For each comparison (CU A– vs. CU A+, MCI A+ and AD A+), we permuted the outcomes (100 iterations) to determine whether brain features were selected by chance. Results for the age-related compositional brain score, the permutated outcome, and the repeated analysis excluding the *APOE* regions are provided in Results SR1.2,1.3,1.4 in supporting information, respectively. All analyses were conducted using R software version 4.2.2.

TABLE 1 Demographic characteristics and *APOE* genotypes distribution in the ALFA sample.

	ALFA *N* = 338	*N*
Sex (*n*, %)		338
Men	127 (37.574%)	
Women	211 (62.426%)	
Age (median, IQR)	57.000 [53.000; 61.000]	338
Education years (median, IQR)	12.000 [11.000; 17.000]	338
Aβ status (*n*, %)		338
Aβ −	220 (65.089%)	
Aβ +	118 (34.911%)	
APOE ε4 carriers (*n*, %)		338
Non-carrier	155 (45.858%)	
Carrier	183 (54.142%)	
APOE ε4 load (*n*, %)		338
Non-carriers	155 (45.858%)	
One ε4 allele	152 (44.970%)	
Two ε4 alleles	31 (9.172%)	
Mini-Mental State Examination (median, IQR)	29.000 [28.000; 30.000]	338

Abbreviations: Aβ, amyloid beta; ALFA, Alzheimer's and Families study; *APOE*, apolipoprotein E; IQR, interquartile range.

3 | RESULTS

3.1 | Characteristics of the sample

The ALFA sample (*N* = 338) included 62.43% women and 54.14% *APOE* ε4 carriers, with a median age of 57 years (interquartile range [IQR] 53–61; Table 1). The ADNI sample (*N* = 330) was defined by 40% women and ≈ 64.55% *APOE* ε4 carriers, with a median age of 74 years (IQR 68.2–78). In the sample there were 30% AD dementia patients and 70% MCI individuals, all of them amyloid positive (Table 2). Differences were assessed among disease-stage groups along the continuum (Table 3). The highest percentage of women was found in ALFA CU A− individuals (63%) while the lowest was observed in ADNI MCI individuals (38%; $p < 0.001$). Significant differences in *APOE* ε4 carriership were observed among groups, with the lowest percentage found in ALFA CU A− participants (40% carriers) and the highest in ALFA CU A+ (78% carriers; $p < 0.001$). Notably, the *APOE* ε4 carrier rate in ALFA CU A+ was higher than in clinical groups from ADNI ($p < 0.001$). There were significant differences in the genetic predisposition to AD between CU A− and CU A+, MCI and AD dementia individuals (FDR *p* value pairwise comparisons < 0.001), but not between CU A+ and the impaired groups. Non-significant differences were found between groups when the *APOE* region was removed from the polygenic risk score (Figure S1 in supporting information). Significant differences were also found in the median value of cortical and subcortical regions between groups (Table S3 in supporting information).

3.2 | Compositional brain scores associated with AD disease stages compared to CU A− individuals

In the ALFA study, the compositional brain score associated with higher odds of belonging to the CU A+ group was primarily characterized by the relative volumetric variation between the pallidum and the amygdala, along with variations in other cortical regions of the parietal, frontal, and temporal lobes as well as subcortical areas (Figure 2). The compositional brain score linked to higher probability of belonging to the MCI A+ group mainly involved the relative volumetric variation between the fusiform and the frontal pole, with additional variations in other temporal and frontal areas, regions of the occipital lobe, and subcortical areas such as the putamen and pallidum (Figure 2). Finally, the compositional brain score associated with higher odds of belonging to the AD A+ group was primarily characterized by the relative volumetric variation between the insula and the inferior parietal, alongside variations in other parietal and frontal regions (Figure 2). Compositional brain scores showed high prediction accuracy in distinguishing CU A− from other disease stages, with robust discrimination for both MCI A+ (area under the curve [AUC] = 0.992) and AD A+ (AUC = 0.980; Table S4, Figure S2 in supporting information). Sensitivity analyses displayed similar compositional brain scores after neuroCombat harmonization (Figures S3–S7, Table S5 in supporting information), particularly for the comparisons between CU A− and both CU A+ and AD A+ participants (Figure S8, Results SR1.1 in supporting information). Post hoc analyses showed that brain regions contributing to the compositional brain scores appeared

TABLE 2 Demographic characteristics and *APOE* genotypes distribution in the ADNI sample.

	ADNI *N* = 330	*N*
Sex (*n*, %)		330
Women	132 (40.000%)	
Men	198 (60.000%)	
Age (median, IQR)	74.050 [68.200; 78.000]	330
Education years (median, IQR)	16.000 [14.000; 18.000]	330
Diagnosis (*n*, %)		330
AD	100 (30.303%)	
MCI	230 (69.697%)	
APOE ε4 carriers (*n*, %)		330
Non-carrier	117 (35.455%)	
Carrier	213 (64.545%)	
APOE ε4 load (*n*, %)		330
Non-carriers	117 (35.455%)	
One ε4 allele	153 (46.364%)	
Two ε4 alleles	60 (18.182%)	
Mini-Mental State Examination (median, IQR)	27.000 [25.000; 29.000]	330

Abbreviations: AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; *APOE*, apolipoprotein E; IQR, interquartile range; MCI, mild cognitive impairment.

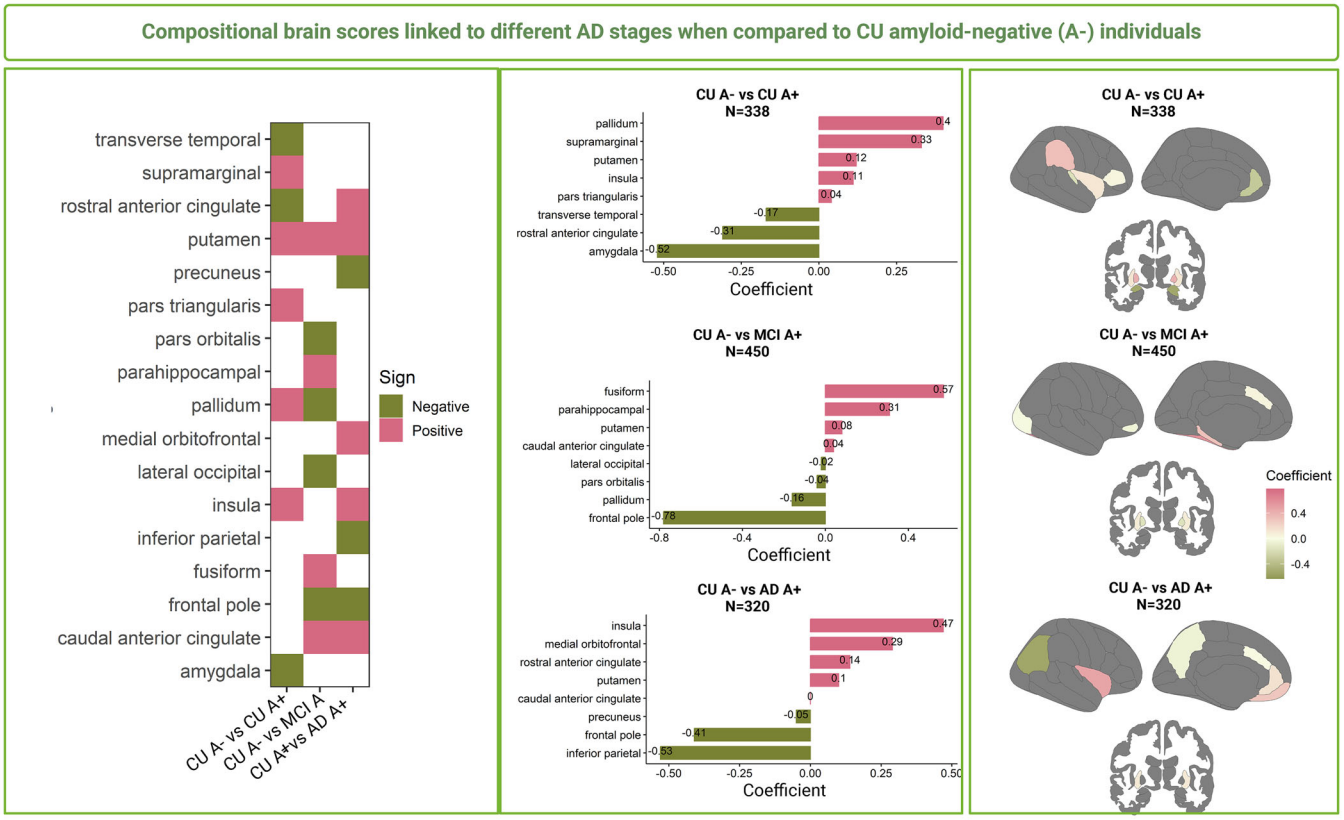

FIGURE 2 Compositional brain scores associated with increased odds of belonging to non-CU A– groups at different AD stages. Regions in dark pink indicate positive contributions to the compositional brain scores, while regions in dark green indicate negative contributions. AD, Alzheimer's disease; AD A+, Alzheimer's disease amyloid-β–positive individuals; CU A–, cognitively unimpaired amyloid-β–negative individuals; CU A+, cognitively unimpaired amyloid-β–positive individuals; MCI A+, mild cognitive impaired amyloid- beta–positive individuals.

TABLE 3 Demographic characteristics and *APOE* genotypes distribution in the ALFA and ADNI sample.

	CU A− ALFA *N* = 220	CU A+ ALFA *N* = 118	MCI A+ ADNI *N* = 230	AD A+ ADNI *N* = 100	*p* value	*p* value CU A− versus CU A+	*p* value CU A− versus MCI	*p* value AD versus CU A−	*p* value CU A+ versus MCI	*p* value AD versus CU A+	*p* value AD versus MCI
Sex (*n*,%)					**<0.001**	0.784	**<0.001**	**0.002**	**<0.001**	**0.018**	0.649
Women	139 (63.182%)	72 (61.017%)	89 (38.696%)	43 (43.000%)							
Men	81 (36.818%)	46 (38.983%)	141 (61.304%)	57 (57.000%)							
Age (median, IQR)	56.000 [52.750; 60.000]	58.000 [54.000; 62.000]	72.850 [68.125; 77.400]	75.600 [69.550; 80.300]	**<0.001**	**0.004**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.025
Education years (median, IQR)	12.500 [11.000; 17.000]	12.000 [10.250; 17.000]	16.000 [14.000; 18.000]	16.000 [14.000; 18.000]	**<0.001**	0.495	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.105
APOE ε4 carriers (*n*,%)					**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.317
Non-carrier	130 (59.091%)	25 (21.186%)	86 (37.391%)	31 (31.000%)							
Carrier	90 (40.909%)	93 (78.814%)	144 (62.609%)	69 (69.000%)							
APOE ε4 load (*n*,%)					**<0.001**	**<0.001**	**<0.001**	**<0.001**	**0.007**	0.170	0.528
Non-carriers	130 (59.091%)	25 (21.186%)	86 (37.391%)	31 (31.000%)							
One ε4 allele	79 (35.909%)	73 (61.864%)	104 (45.217%)	49 (49.000%)							
Two ε4 alleles	11 (5.000%)	20 (16.949%)	40 (17.391%)	20 (20.000%)							
PRS-AD (median, IQR)	−0.038 [−0.155; 0.090]	0.056 [−0.086; 0.220]	0.033 [−0.112; 0.204]	0.070 [−0.054; 0.234]	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.559	0.459	0.150
PRS−ADno*APOE* (median, IQR)	−0.11 [−0.18; −0.02]	−0.12 [−0.22; 0.03]	−0.10 [−0.19; −0.01]	−0.08 [−0.19; 0.02]	0.605	0.903	0.708	0.708	0.921	0.708	0.708
Mini-Mental State Examination (median, IQR)	29.000 [28.000; 30.000]	29.000 [29.000; 30.000]	28.000 [27.000; 29.000]	23.000 [21.750; 25.000]	**<0.001**	0.303	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**

Note: Values highlighted in bold font indicate an FDR-adjusted *p*-value < 0.05 for the pairwise comparisons.

Abbreviations: AD, Alzheimer's Disease; AD A+, Alzheimer's disease amyloid-β–positive individuals; ADNI, Alzheimer's Disease Neuroimaging Initiative; ALFA, Alzheimer's and Families study; *APOE*, apolipoprotein E; CU A−, cognitively unimpaired amyloid-β–negative individuals; CU A+, cognitively unimpaired amyloid-β–positive individuals; MCI A+, mild cognitive impaired amyloid-β–positive individuals; PRS_{AD}, polygenic risk score of AD.

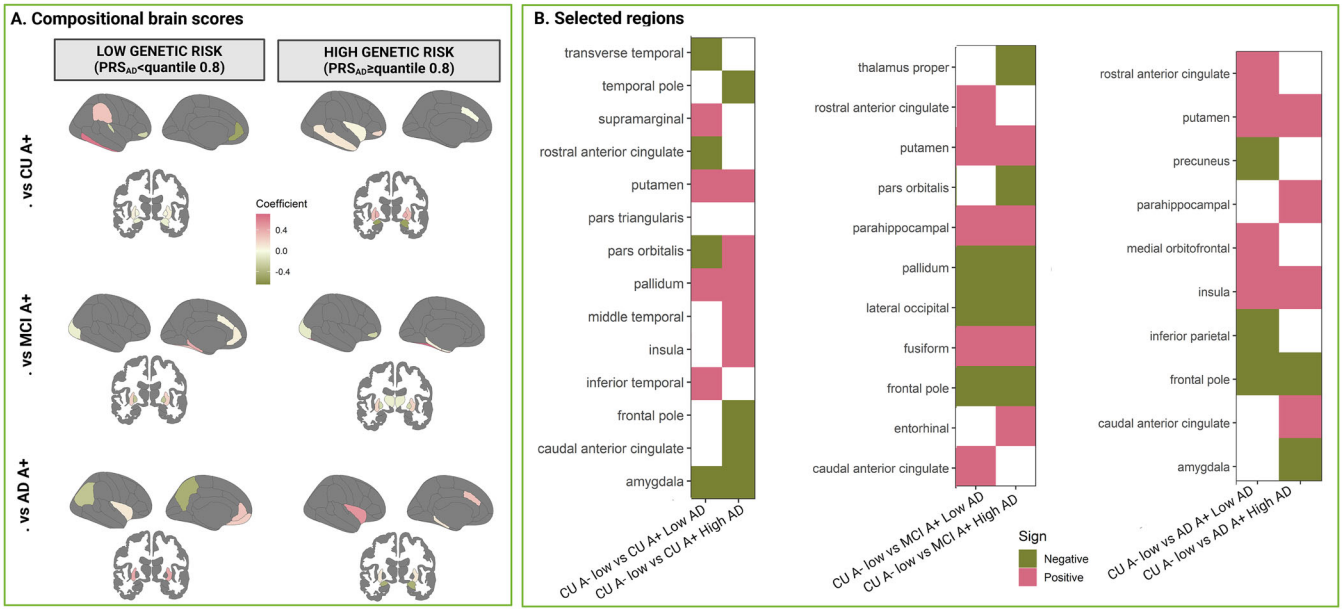

FIGURE 3 Compositional brain scores associated with increased odds of belonging to the non-CU A– group stratifying by AD risk profile at different AD stages. Regions in dark pink indicate positive contributions to the compositional brain scores, while regions in dark green indicate negative contributions. AD, Alzheimer's disease; AD A+, Alzheimer's disease amyloid-β-–positive individuals; CU A–, cognitively unimpaired amyloid-β–negative individuals; CU A+, cognitively unimpaired amyloid-β–positive individuals; MCI A+, mild cognitive impaired amyloid-β–positive individuals.

stable and were less commonly included in the main brain scores when the outcome was permuted (Results SR1.2, Figure S9 in supporting information). Moreover, the compositional brain scores associated with higher odds of belonging to the non-CU A– group along the disease continuum involved different regions than those associated with higher chronological age in the CU A– group (Results SR1.3, Figures S10–S11 in supporting information).

3.3 | AD genetic risk influences brain structure at each stage of the disease continuum

When stratified by AD genetic risk profile, the compositional brain score associated with higher odds of belonging to the CU A+ group remained largely unchanged for the CU A+ participants at low risk of AD (Figure 3). The putamen, pallidum, and amygdala consistently contributed to the compositional brain score, regardless of the CU A+ genetic risk profile. Nonetheless, the brain score shifted when focusing on the CU A+ group at high risk of AD. For instance, the temporal and frontal poles, caudal anterior cingulate, pars orbitalis, and middle temporal regions showed vulnerability to AD genetic risk in CU A+.

Minor differences were observed in the compositional brain score associated with higher odds of belonging to the MCI A+ group when split by AD genetic risk (Figure 3). The entorhinal and thalamus contributed to the relative volumetric variations linked to a higher probability of belonging to the high-risk MCI A+ group. The rostral anterior cingulate was uniquely involved in the brain compositional score with higher odds of belonging to the low-risk MCI A+ group. Nonetheless, the contributions of the putamen, parahippocampal gyrus, fusiform, together with the palladium, lateral occipital, and frontal pole to the brain score associated with higher odds of belonging to the MCI A+ group, were independent of the AD genetic burden. Last, compositional brain scores associated with a higher probability of being AD A+ varied slightly across genetic risk-stratified models (Figure 3). The amygdala and parahippocampal regions were particularly vulnerable to genetic risk and were linked to a higher likelihood of belonging to the high-risk AD A+ group. However, the contributions of the putamen, insula, and frontal pole to the brain score related to higher odds of belonging to the AD A+ group, were unaffected by the AD genetic burden of the participants. Compositional brain scores showed high prediction accuracy in distinguishing at low genetic risk CU A– from both at low and at high genetic risk MCI A+ and AD A+ (Table S6, Figure S12 in supporting information). The same analyses were conducted using a PRS that excluded the *APOE* region (PRS$_{ADnoAPOE}$; Results SR1.4–1.5, Figures S13–S15, and Table S7 in supporting information).

4 | DISCUSSION

This study leverages CoDA to analyze neuroimaging data, uncovering structural brain variations across AD stages and assessing genetic risk impacts. CoDA identified compositional brain scores that differentiate disease stages along the AD continuum and reveal how genetic risk modulates specific stage-related patterns. These findings highlight CoDA's unique ability to capture complex interdependencies among brain regions often missed by traditional methods, offering a complimentary framework for structural brain analysis and new insights into AD distinct patterns.

Building on previous neuroimaging research on AD-related brain changes,[42] this study identifies stage-specific compositional brain scores, with the amygdala and pallidum, despite being inversely related, contributing most significantly to the CU A+ group score. Previous studies showed early divergence of structures like the amygdala and hippocampus in AD pathology, with the amygdala showing proportionally greater changes.[6,43] Amygdala atrophy has also been linked to cognitive decline, predicting Mini-Mental State Examination scores and hippocampal atrophy in mild AD.[44] Overall, our findings suggest an early impact of amyloid pathology on brain structure in CU individuals, notably targeting the amygdala, an early hallmark of AD.[45–47] Frontal regions and the anterior cingulate also contributed to the CU A+ brain score, possibly indicating resilience-related mechanisms in early structural changes.[48]

Moreover, the fusiform and frontal pole, despite their inverse relationship, contributed most significantly to the compositional brain score associated with higher odds of MCI A+ group. Recent studies showed distinct atrophy patterns in MCI individuals, with fusiform and frontal thinning observed in different MCI subtypes,[49] highlighting the complexity of structural changes across profiles. The compositional score also involved regions related to memory and executive function, indicating more pronounced changes as resilience declines and memory and executive impairments emerge.[50]

Finally, the insula and inferior parietal regions, though inversely related, contributed the most to the compositional score linked to AD A+ likelihood. Additional contributions came from cortical areas in the frontal and parietal lobes, the insula, cingulate cortex, and subcortical areas such as the putamen. Previous studies in CU, MCI, and AD patients showed decreased nodal centrality in the middle temporal and increased centrality in the precuneus (parietal) in MCI and AD compared to CU.[42] Overall, the compositional score involved regions linked to memory, executive, and motor functions, suggesting compensatory mechanisms in preserved areas (frontal and parietal) to offset losses in affected regions like the temporal lobe and amygdala.

Stratifying by AD genetic risk profile (high/low) within each disease stage (CU A+, MCI A+, and AD A+), we identified genetically vulnerable regions linked to higher odds of belonging to non-CU A− compared to high-risk individuals at each stage of the disease. For instance, the temporal and frontal poles, caudal anterior cingulate, pars orbitalis, and middle temporal were genetically vulnerable in CU A+. The entorhinal and thalamus proper exhibited genetic vulnerability in MCI A+, whereas the amygdala and parahippocampal regions were vulnerable in AD A+. Notably, the amygdala also contributed to the CU A+ compositional brain score, suggesting a dual role across the disease continuum.[43] Results suggest that, early on, amyloid pathology may primarily affect the amygdala, while genetic factors accelerate its atrophy in later AD stages.[6,51]

The effect of *APOE* on brain structure remained largely consistent when genetic groups were defined using an AD PRS excluding *APOE*-related genetic variants. Brain scores for CU A− versus MCI A+ and AD A+ were nearly identical, though minor differences appeared in genetic-stratified models for CU A− versus CU A+ particularly in the preclinical stage model, when CU A+ were younger and had a lower proportion of A+T+ subjects. This suggests an earlier impact of *APOE*-related variants on structural brain changes. Recent studies revealed distinct genetic architectures for younger versus older onset, with chromosome 19 variants explaining half of the heritability in the younger group, highlighting age as a modifier of *APOE*'s AD risk.[52] Other studies showed that *APOE* ε4 carriership was linked to A−T− to A+T− conversion, while AD polygenic risk drove progression to A+T+,[53] suggesting *APOE*'s stronger early influence on brain structure, likely via amyloid pathways. In contrast, for MCI A+ and AD A+ stages, compositional brain scores differed significantly among genetic risk, independent of *APOE*. These findings indicate that in later stages, overall AD genetic risk, beyond amyloid, may influence pathways such as tau, immunity, and lipid processing.[54–56]

Last, we found that the putamen, frontal pole, and insula contributed non-specifically to structural changes across AD stages, overlapping with regions in the age-related compositional brain score. The putamen consistently showed a positive contribution across all brain scores, where its relative increase, along with other regional changes, was linked to higher likelihood of non-belonging to the CU A− group. However, its relative decrease, together with volumetric changes in other regions, was associated with higher chronological age. Prior studies revealed a linear decline in putamen volume with age,[57–59] suggesting its role as an aging marker. Given age's major influence on AD risk, putamen involvement across compositional brain scores likely reflects the impact of aging along the AD continuum.

Several regions overlapped across pairwise compositional brain scores. The caudal anterior cingulate and frontal pole were linked to a higher likelihood of belonging to the MCI A+ and AD A+ groups, while the insula and rostral anterior cingulate contributed to the CU A+ and AD A+ compositional brain scores. Furthermore, the pallidum contributed to CU A+ and MCI A+ compositional brain scores. These findings aligned with recent machine learning (ML) studies which identified temporal regions, specifically the hippocampus, cingulate, and frontal and parietal lobes, as key for distinguishing between CU and AD individuals.[16,17] While ML highlighted critical morphological features, it did not capture the relative changes between these regions.

Finally, we validated the robustness of the variable selection procedure by permuting the outcome across all pairwise models. This approach confirmed minimal or no overlap between regions in the permuted compositional brain scores and those identified in the main compositional brain scores from comparisons between CU A− and CU A+, MCI A+ and AD A+. This strong distinction underscores the reliability of our findings and supports specificity of the identified brain regions to each disease stage.

This study showed that CoDA emerged as an alternative method for addressing limitations in traditional neuroimaging analyses, providing an accurate and comprehensive understanding of brain structural alterations.[60] Unlike conventional univariate or multivariate analyses, which capture correlated features[61] but often miss the compositional nature of brain data, thereby risking spurious correlations, CoDA effectively models complex interdependencies among brain regions by focusing on relative variations. In addition, a notable strength of CoDA is its ability to integrate neuroimaging data from two independent

cohorts without rigorous harmonization requirements, overcoming a key barrier in traditional analyses.[39] This flexibility enables the analysis of diverse datasets, improving cross-study reliability through standardization (via a log-ratio approach), even when labeling conventions differ. Such robustness supports collaborative, large-scale research efforts.

Despite its strengths, we acknowledge the need for longitudinal data to explore the trajectory of the volumetric variations of the brain regions over time. Additionally, the use of a binary PRS to assess genetic impact on brain structure also introduces limitations. While dichotomizing PRS provides a general AD risk profile, it may oversimplify complex interactions among genetics, age, and disease stage. Moreover, future studies should address interactions between genetics and environmental factors[62] not covered in this study.

In summary, this study identified distinct brain structure patterns associated with AD and showed the influence of higher AD genetic risk on interdependent changes across brain regions. Using CoDA, we captured AD-specific brain patterns that diverge from those seen in normal aging, underscoring CoDA's potential as a robust multivariate framework for exploring AD-related structural changes. By examining genetic risk–stratified models, the findings revealed a broader genetic landscape affecting brain morphology in AD, beyond the well-known *APOE ε4* allele, thus highlighting the complex genetic interplay that drives brain structure alterations. This insight is particularly valuable for understanding AD pathogenesis and paves the way for targeted interventions tailored to individuals at higher genetic risk, supporting both research advancements and clinical applications to improve brain health.

ACKNOWLEDGMENTS

This publication is part of the ALFA study. The authors express their most sincere gratitude to the ALFA project participants and relatives without whom this research would have not been possible. Collaborators of the ALFA study are: Müge Akinci, Federica Anastasi, Annabella Beteta, Raffaele Cacciaglia, Lidia Canals, Alba Cañas, Carme Deulofeu, Maria Emilio, Irene Cumplido-Mayoral, Marta del Campo, Carme Deulofeu, Ruth Dominguez, Maria Emilio, Karine Fauria, Sherezade Fuentes, Marina García, Laura Hernández, Gema Huesa, Jordi Huguet, Laura Iglesias, Esther Jiménez, David López-Martos, Paula Marne, Tania Menchón, Paula Ortiz-Romero, Eleni Palpatzis, Wiesje Pelkmans, Albina Polo, Sandra Pradas, Mahnaz Shekari, Lluís Solsona, Anna Soteras, Núria Tort-Colet, and Marc Vilanova. Data collection and sharing for this project were funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. The research leading to these results has received funding from "la Caixa" Foundation (ID 100010434), under agreement LCF/PR/GN17/50300004, the Health Department of the Catalan Government (Health Research and Innovation Strategic Plan (PERIS) 2016-2020 grant# SLT002/16/00201), and the Alzheimer's Association and an international anonymous charity foundation through the TriBEKa Imaging Platform project (TriBEKa-17-519007). Additional support has been received from the Universities and Research Secretariat, Ministry of Business and Knowledge of the Catalan Government under the grant no. 2021 SGR 00913. All CRG authors acknowledge the support of the Spanish Ministry of Science, Innovation, and Universities to the EMBL partnership, the Centro de Excelencia Severo Ochoa, and the CERCA Programme/Generalitat de Catalunya. N.V.-T. was supported by the Spanish Ministry of Science and Innovation—State Research Agency (IJC2020-043216-I/MCIN/AEI/10.13039/501100011033) and the European Union «NextGenerationEU»/PRTR and currently receives funding from the Spanish Research Agency MICIU/AEI/10.13039/501100011033 (grant RYC2022-038136-I cofunded by the European Union FSE+ and grant PID2022-143106OA-I00 cofunded by the European Union FEDER). In addition, N.V.-T. is supported by the William H. Gates Sr. Fellowship from the Alzheimer's Disease Data Initiative. Data partially used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

CONFLICT OF INTEREST STATEMENT

J.D.G. has served as a consultant for Roche Diagnostics and Prothena Biosciences; he has given lectures at symposiums sponsored by General Electric, Philips, Esteve, Life-MI, and Biogen; and he received research support from GE Healthcare, Roche Diagnostics, and Hoffmann-La Roche. The remaining co-authors have no conflicts to disclose. Author disclosures are available in the supporting information.

CONSENT STATEMENT

The ALFA study was conducted in accordance with the directives of the Spanish Law 14/2007, of July 3, on Biomedical Research (Ley 14/2007 de Investigación Biomédica). The ALFA study protocol was approved by the Independent Ethics Committee Parc de Salut Mar Barcelona and registered at ClinicalTrials.gov (Identifier: NCT01835717). All participants accepted the study procedures by signing the study's informed consent form that had also been approved by the same institutional review board.

ORCID

Natalia Vilor-Tejedor https://orcid.org/0000-0003-4935-6721

REFERENCES

1. Andrade-Guerrero J, Santiago-Balmaseda A, Jeronimo-Aguilar P, et al. Alzheimer's disease: an updated overview of its genetics. *Int J Mol Sci*. 2023;24:3754. doi:10.3390/ijms24043754
2. Avelar-Pereira B, Belloy ME, O'Hara R, Hosseini SMH. Decoding the heterogeneity of Alzheimer's disease diagnosis and progression using multilayer networks. *Mol Psychiatry*. 2022;28:2423.doi:10.1038/s41380-022-01886-z
3. Jack CR Jr, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimers Dement*. 2024;20:5143-5169. doi:10.1002/alz.13859
4. Thompson PM, Hayashi KM, Dutton RA, et al. Tracking Alzheimer's disease. *Ann N Y Acad Sci*. 2007;1097:183-214. doi:10.1196/annals.1379.017
5. Apostolova LG, Thompson PM. Mapping progressive brain structural changes in early Alzheimer's disease and mild cognitive impairment. *Neuropsychologia*. 2008;46:1597-1612. doi:10.1016/j.neuropsychologia.2007.10.026
6. Planche V, Manjon JV, Mansencal B, et al. Structural progression of Alzheimer's disease over decades: the MRI staging scheme. *Brain Commun*. 2022;4:fcac109. doi:10.1093/braincomms/fcac109
7. Zhao B, Ibrahim JG, Li Y, et al. Heritability of regional brain volumes in large-scale neuroimaging and genetic studies. *Cereb Cortex*. 2019;29:2904-2914. doi:10.1093/cercor/bhy157
8. Vilor-Tejedor N, Evans TE, Adams HH, et al. Genetic influences on hippocampal subfields: an emerging area of neuroscience research. *Neurol Genet*. 2021;7:e591. doi:10.1212/NXG.0000000000000591
9. Wang H, Yang F, Zhang S, Xin R, Sun Y. Genetic and environmental factors in Alzheimer's and Parkinson's diseases and promising therapeutic intervention via fecal microbiota transplantation. *Npj Parkinson's Disease*. 2021;7:1-10. doi:10.1038/s41531-021-00213-7
10. Zhao Y-L, Qu Y, Ou Y-N, Zhang Y-R, Tan L, Yu J-T. Environmental factors and risks of cognitive impairment and dementia: a systematic review and meta-analysis. *Ageing Res Rev*. 2021;72:101504. doi:10.1016/j.arr.2021.101504
11. Kannappan B, Gunasekaran TI, Te Nijenhuis J, et al. Alzheimer's Disease Neuroimaging Initiative. Polygenic score for Alzheimer's disease identifies differential atrophy in hippocampal subfield volumes. *PLoS One*. 2022;17:e0270795. doi:10.1371/journal.pone.0270795
12. Foo H, Thalamuthu A, Jiang J, et al. Associations between Alzheimer's disease polygenic risk scores and hippocampal subfield volumes in 17,161 UK Biobank participants. *Neurobiol Aging*. 2021;98:108-115. doi:10.1016/j.neurobiolaging.2020.11.002
13. Murray AN, Chandler HL, Lancaster TM. Multimodal hippocampal and amygdala subfield volumetry in polygenic risk for Alzheimer's disease. *Neurobiol Aging*. 2021;98:33-41. doi:10.1016/j.neurobiolaging.2020.08.022
14. Sheng J, Wang L, Cheng H, Zhang Q, Zhou R, Shi Y. Strategies for multivariate analyses of imaging genetics study in Alzheimer's disease. *Neurosci Lett*. 2021;762:136147. doi:10.1016/j.neulet.2021.136147
15. Vilor-Tejedor N, Garrido-Martín D, Rodriguez-Fernandez B, Lamballais S, Guigó R, Gispert JD. Multivariate analysis and modelling of multiple brain endOphenotypes: let's MAMBO!. *Comput Struct Biotechnol J*. 2021;19:5800-5810. doi:10.1016/j.csbj.2021.10.019
16. Diogo VS, Ferreira HA, Prata D.; Alzheimer's Disease Neuroimaging Initiative. Early diagnosis of Alzheimer's disease using machine learning: a multi-diagnostic, generalizable approach. *Alzheimers Res Ther*. 2022;14:107. doi:
17. Donnelly-Kehoe PA, Pascariello GO, Gómez JC., & Alzheimer's Disease Neuroimaging Initiative. Looking for Alzheimer's disease morphometric signatures using machine learning techniques. *J Neurosci Methods*. 2018;302:24-34. doi:10.1016/j.jneumeth.2017.11.013
18. Xu L, Groth KM, Pearlson G, Schretlen DJ, Calhoun VD. Source-based morphometry: the use of independent component analysis to identify gray matter differences with application to schizophrenia. *Hum Brain Mapp*. 2009;30:711-724. doi:10.1002/hbm.20540
19. Luz Calle M. Statistical analysis of metagenomics data. *Genomics Inform*. 2019;17:e6. doi:10.5808/GI.2019.17.1.e6
20. Eickhoff SB, Yeo BTT, Genon S. Imaging-based parcellations of the human brain. *Nat Rev Neurosci*. 2018;19:672-686. doi:10.1038/s41583-018-0071-7
21. Lawrence RM, Bridgeford EW, Myers PE, et al. Standardizing human brain parcellations. *Sci Data*. 2021;8:78. doi:10.1038/s41597-021-00849-3
22. Calle ML, Pujolassos M, Susin A. coda4microbiome: compositional data analysis for microbiome cross-sectional and longitudinal studies. *BMC Bioinformatics*. 2023;24:82. doi:10.1186/s12859-023-05205-3
23. Wang B, Caffo BS, Luo X, et al., & Alzheimer's Disease Neuroimaging Initiative. Regularized regression on compositional trees with application to MRI analysis. *J R Stat Soc Ser C Appl Stat*. 2022;71:541-561. doi:10.1111/rssc.12545
24. Zhao Y, Wang B, Liu C-F, et al. Identifying brain hierarchical structures associated with Alzheimer's disease using a regularized regression method with tree predictors. *Biometrics*. 2023;79:2333-2345. doi:10.1111/biom.13775
25. Bueno-Notivol J, Gracia-García P, Olaya B, de la Cámara C, López-Antón R, Santabárbara J. An Alzheimer's dementia cumulative risk model in a sample of general population over 65: public health implications. *Eur J Psychiatry*. 2022;37(2):117-124. doi:10.1016/j.ejpsy.2022.09.006
26. Molinuevo JL, Gramunt N, Gispert JD, et al. The ALFA project: a research platform to identify early pathophysiological features of Alzheimer's disease. *Alzheimers Dement*. 2016;2:82-92. doi:10.1016/j.trci.2016.02.003
27. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology*. 2010;74:201-209. doi:10.1212/WNL.0b013e3181cb3e25
28. Jack CR Jr, Bennett DA, Blennow K, et al. A/T/N: an unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology*. 2016;87:539-547. doi:10.1212/WNL.0000000000002923
29. Milà-Alomà M, Salvadó G, Gispert JD, et al., & ALFA study. Amyloid beta, tau, synaptic, neurodegeneration, and glial biomarkers in the preclinical stage of the Alzheimer's continuum. *Alzheimers Dement*. 2020;16:1358-1371. doi:10.1002/alz.12131
30. Blauwendraat C, Faghri F, Pihlstrom L, et al. NeuroChip, an updated version of the NeuroX genotyping platform to rapidly screen for variants associated with neurological diseases. *Neurobiol Aging*. 2017;57:247.e9-e13. doi:10.1016/j.neurobiolaging.2017.05.009
31. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48:1284-1287. doi:10.1038/ng.3656
32. Vilor-Tejedor N, Genius P, Rodríguez-Fernández B, et al.. Genetic characterization of the ALFA study: uncovering genetic profiles in the Alzheimer's *continuum*. *Alzheimers Dement*. 2023. doi:10.1002/alz.13537

33. Weber CJ, Carrillo MC, Jagust W, et al. The Worldwide Alzheimer's Disease Neuroimaging Initiative: aDNI-3 updates and global perspectives. *Alzheimers Dement*. 2021;7:e12226. doi:10.1002/trc2.12226
34. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience*. 2019;8:giz082. doi:10.1093/gigascience/giz082
35. Wightman DP, Jansen IE, Savage JE, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat Genet*. 2021;53:1276-1282. doi:10.1038/s41588-021-00921-z
36. Huguet J, Falcon C, Fusté D, et al., & ALFA study. Management and quality control of large neuroimaging datasets: developments from the Barcelonaβeta Brain Research Center. *Front Neurosci*. 2021;15:633438. doi:10.3389/fnins.2021.633438
37. Jack CR Jr, Bernstein MA, Fox NC, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging*. 2008;27:685-691. doi:10.1002/jmri.21049
38. Fortin J-P, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*. 2018;167:104-120. doi:10.1016/j.neuroimage.2017.11.024
39. Richter S, Winzeck S, Correia MM, et al. Validation of cross-sectional and longitudinal ComBat harmonization methods for magnetic resonance imaging data on a travelling subject cohort. *Neuroimage Rep*. 2022;2:None. doi:10.1016/j.ynirp.2022.100136
40. Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006;31:968-980. doi:10.1016/j.neuroimage.2006.01.021
41. Susin A, Wang Y, Lê Cao K-A, Calle ML. Variable selection in microbiome compositional data analysis. *NAR Genom Bioinform*. 2020;2:lqaa029. doi:10.1093/nargab/lqaa029
42. Yao Z, Zhang Y, Lin L, Zhou Y, Xu C, Jiang T.,& Alzheimer's Disease Neuroimaging Initiative. Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease. *PLoS Comput Biol*. 2010;6:e1001006. doi:10.1371/journal.pcbi.1001006
43. Coupé P, Manjón JV, Lanuza E, Catheline G. Lifespan changes of the human brain in Alzheimer's disease. *Sci Rep*. 2019;9:3998. doi:10.1038/s41598-019-39809-8
44. Poulin SP, Dautoff R, Morris JC, Barrett LF, Dickerson BC, &Alzheimer's Disease Neuroimaging Initiative. Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. *Psychiatry Res*. 2011;194:7-13. doi:10.1016/j.pscychresns.2011.06.014
45. Knafo S, Venero C, Merino-Serrais P, et al. Morphological alterations to neurons of the amygdala and impaired fear conditioning in a transgenic mouse model of Alzheimer's disease. *J Pathol*. 2009;219:41-51. doi:10.1002/path.2565
46. Folorunso OO, Harvey TL, Brown SE, Chelini G, Berretta S, Balu DT. The D-serine biosynthetic enzyme serine racemase is expressed by reactive astrocytes in the amygdala of human and a mouse model of Alzheimer's disease. *Neurosci Lett*. 2023;792:136958. doi:10.1016/j.neulet.2022.136958
47. Al-Ani L, Tao A, Dyke J, Chiang G, Ishii M. Amygdala atrophy as an early manifestation of Alzheimer's Disease (S39.005). *Neurology*. 2023;100:1800. doi:10.1212/wnl.0000000000202128
48. Barulli D, Stern Y. Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. *Trends Cogn Sci*. 2013;17:502-509. doi:10.1016/j.tics.2013.08.012
49. Machulda MM, Lundt ES, Albertson SM, et al. Cortical atrophy patterns of incident MCI subtypes in the Mayo Clinic Study of Aging. *Alzheimers Dement*. 2020;16:1013-1022. doi:10.1002/alz.12108
50. Small SA, Duff K. Linking Abeta and tau in late-onset Alzheimer's disease: a dual pathway hypothesis. *Neuron*. 2008;60:534-542. doi:10.1016/j.neuron.2008.11.007
51. Lupton MK, Strike L, Hansell NK, et al. The effect of increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume. *Neurobiol Aging*. 2016;40:68-77. doi:I:10.1016/j.neurobiolaging.2015.12.023
52. Lo MT, Kauppi K, Fan CC, et al., & Alzheimer's Disease Genetics Consortium. Identification of genetic heterogeneity of Alzheimer's disease across age. *Neurobiol Aging*. 2019;84:243.e1-e9. doi:10.1016/j.neurobiolaging.2019.02.022
53. Mattsson-Carlgren N. Disentangling genetic risks for development and progression of Alzheimer's disease. *Brain*. 2024;147:2604-2606. doi:10.1093/brain/awae237
54. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat Genet*. 2019;51:414-430. doi:10.1038/s41588-019-0358-2
55. Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet*. 2022;54:412-436. doi:10.1038/s41588-022-01024-z
56. Lorenzini L, Collij LE, Tesi N, et al. Alzheimer's disease genetic pathways impact cerebrospinal fluid biomarkers and imaging endophenotypes in non-demented individuals. *Alzheimers Dement*. 2024;20(9):6146-6160. doi:10.1002/alz.14096
57. Zheng F, Liu Y, Yuan Z, et al. Age-related changes in cortical and subcortical structures of healthy adult brains: a surface-based morphometry study. *J Magn Reson Imaging*. 2019;49:152-163. doi:10.1002/jmri.26037
58. Ostby Y, Tamnes CK, Fjell AM, Westlye LT, Due-Tønnessen P, Walhovd KB. Heterogeneity in subcortical brain development: a structural magnetic resonance imaging study of brain maturation from 8 to 30 years. *J Neurosci*. 2009;29:11772-11782. doi:10.1523/JNEUROSCI.1242-09.2009
59. Pfefferbaum A, Rohlfing T, Rosenbloom MJ, Chu W, Colrain IM, Sullivan EV. Variation in longitudinal trajectories of regional brain volumes of healthy men and women (ages 10 to 85 years) measured with atlas-based parcellation of MRI. *Neuroimage*. 2013;65:176-193. doi:10.1016/j.neuroimage.2012.10.008
60. Xin Y, Sheng J, Miao M, Wang L, Yang Z. A review of imaging genetics in Alzheimer's disease. *Journal of Clinical*. 2022. doi:10.1016/j.jocn.2022.04.017
61. Cruciani F, Altmann A, Lorenzi M, Menegaz G, Galazzo IB. What PLS can still do for Imaging Genetics in Alzheimer's disease. In: 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), *IEEE*; 2022, p. 1-4. doi:0.1109/BHI56158.2022.9926813
62. Wang C, Sun J, Guillaume B, Ge T, et al. A set-based mixed effect model for gene-environment interaction and its application to neuroimaging phenotypes. *Front Neurosci*. 2017;11:191. doi:10.3389/fnins.2017.00191

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Genius P, Calle ML, Rodríguez-Fernández B, et al. Compositional brain scores capture Alzheimer's disease–specific structural brain patterns along the disease continuum. *Alzheimer's Dement*. 2025;21:e14490. https://doi.org/10.1002/alz.14490