Archival Report

Biological Psychiatry: CNNI

Structural Brain Differences in the Alzheimer's Disease Continuum: Insights Into the Heterogeneity From a Large Multisite Neuroimaging Consortium

Tavia E. Evans, Natalia Vilor-Tejedor, Gregory Operto, Carles Falcon, Albert Hofman, Agustin Ibáñez, Sudha Seshadari, Louis C.S. Tan, Michael Weiner, Suverna Alladi, Udunna Anazodo, Juan Domingo Gispert, for the Alzheimer's Disease Neuroimaging Initiative, for the Australian Imaging Biomarkers and Lifestyle flagship study of ageing, and Hieab H.H. Adams

ABSTRACT

BACKGROUND: Neurodegenerative diseases require collaborative, multisite research to comprehensively grasp their complex and diverse pathological progression; however, there is caution in aggregating global data due to data heterogeneity. In the current study, we investigated brain structure across stages of Alzheimer's disease (AD) and how relationships vary across sources of heterogeneity.

METHODS: Using 6 international datasets ($N > 27{,}000$), associations of structural neuroimaging markers were investigated in relation to the AD continuum via meta-analysis. We investigated whether associations varied across elements of magnetic resonance imaging acquisition, study design, and populations.

RESULTS: Modest differences in associations were found depending on how data were acquired; however, patterns were similar. Preliminary results suggested that neuroimaging marker–AD relationships differ across ethnic groups.

CONCLUSIONS: Diversity in data offers unique insights into the neural substrate of AD; however, harmonized processing and transparency of data collection are needed. Global collaborations should embrace the inherent heterogeneity that exists in the data and quantify its contribution to research findings at the meta-analytical stage.

https://doi.org/10.1016/j.bpsc.2024.07.019

Most neurodegenerative disorders, such as Alzheimer's disease (AD), are highly complex and heterogeneous and are caused by an interplay of many genetic and nongenetic factors, such as the *APOE* gene and smoking. These factors explain the liability to disease within populations and disparities across populations (1,2). However, the effects of each risk factor are generally small, which makes them difficult to detect in typically underpowered neuroimaging studies, which in turn contributes to the alarming lack of reproducibility within the field (3–5). This is more accentuated in nonstereotypical samples (in terms of demographics, ancestral admixture, geographical regions), which intrinsically defy any systematic generalization (6). Multisite efforts are needed to achieve the required sample sizes, but data are difficult to pool together given the large amount of heterogeneity in the acquisition and analysis of neuroimaging data. Impressive initiatives, such as the ADNI (Alzheimer's Disease Neuroimaging Initiative), try to solve this by prospectively collecting data within a harmonized acquisition protocol (7). While valuable, such initiatives are extremely costly and time consuming and therefore not scalable to the sample sizes necessary to study the multifactorial origins of neurodegenerative diseases. Furthermore, underrepresented regions such as Africa, South America, and Asia typically have access to different infrastructure, making it unrealistic that such harmonized data collection will be feasible in the short term (8). An alternative approach is to accept the existing measurement heterogeneity and account for it analytically, which is the basis for the UNITED (Uncovering Neurodegenerative Insights Through Ethnic Diversity) consortium (9). While this heterogeneity contributes to increased noise, it also has several benefits that could potentially outweigh this. First, such inclusive collaborations could drastically increase sample size and thus also statistical power. Imaging genetics consortia have shown that, despite widely varying assessments, replicable findings can be obtained (10–12). Second, finding the same associations across sites with heterogeneous data collection is supportive of a true underlying signal, whereas harmonized data collection could underlie a systematic bias toward scanner and sequence choice that renders the results less generalizable. Third, this approach allows for the inclusion of currently underrepresented groups, which leads to more generalizable results as

ISSN: 2451-9022

well as incorporation of researchers from these regions in global studies. Furthermore, this large, diverse approach may have more potential to capture the biological heterogeneity of the disease. However, the impact of the various layers of measurement heterogeneity in neuroimaging research has not yet been systematically studied in a multisite setting.

In the current investigation, we studied more than 27,000 individuals to investigate associations of structural neuroimaging markers in groups with AD and mild cognitive impairment (MCI) in comparison to a group of healthy control participants (HCs), between individuals with and without a family history of dementia, and between *APOE* ε4 carriers versus noncarriers. We assessed whether these relationships varied across elements of magnetic resonance imaging (MRI) acquisition; study design, such as processing and analysis methods; and population groups, such as ethnicity, so that these insights can be employed for future large collaborative neuroimaging efforts.

METHODS AND MATERIALS

A schematic overview of the study design is shown in Figure 1. In short, we included 27,764 individuals from the UKBB (UK Biobank) (13), ALFA (ALzheimer and FAmilies study) (14), ADNI (7), NACC (National Alzheimer's Coordinating Center) (15), AIBL (Australian Imaging Biomarkers and Lifestyle study) (16), and WMH-AD (White Matter Hyperintensities in Alzheimer's Disease study) (17). All datasets had institutional review board approval and data sharing consent in place, and except for ALFA, all are open access datasets. An additional reproducibility set, ID1000 from the AOMIC (Amsterdam Open MRI Collection), was used to determine the reproducibility of the image processing methods (18). This study included cognitively healthy individuals who underwent a short-time test-retest imaging protocol undertaking that results in 3 T1-weighted images from 1 session. From this sample, 2725 T1 MRI scans from 927 individuals were included. Methods related to clinical and imaging acquisition and processing are described in Table S1 in Supplement 2. Briefly, the diagnoses were predominantly clinical/research diagnoses and not biomarker-based diagnoses, with those classified as AD and MCI having the diagnosis at MRI date. Incident AD indicates a conversion to AD; these individuals were free of AD at the time of the MRI date (healthy control or MCI) and received the classification during follow-up. Descriptive information about these studies, including population characteristics, is presented in Table S2 in Supplement 2.

MRI Processing

T1-weighted images were processed to extract volumetric information using 2 different software packages, FreeSurfer (version 6.0) (19) and FSL (version 6.0.5.1) (20); these variables included the total gray matter, total white matter, and 7 subcortical structures. These variables were standardized within each sample. Next, 54,000 features of the shape of these 7 subcortical brain structures (Figure S1 in Supplement 1) were derived using the ENIGMA-Shape protocol (21), which used the FreeSurfer segmentations as input. Lastly, the T1-weighted images were processed for a voxel-based morphometry (VBM) analysis using a previous pipeline (22), which provides highly detailed information on the structure of the whole brain using approximately 1.5 million variables that are able to be meta-analyzed easily at the voxel level. The full methodology is reported in Supplemental Methods in Supplement 1.

Statistical Analysis

Briefly, first the software FSL, FreeSurfer, ENIGMA-Shape, and VBM were analyzed to examine the reproducibility of processing methods through an intraclass correlation coefficient (ICC) analysis of repeat imaging. For vertex and voxel analysis, summary measures for each structure or the total gray matter (VBM) were reported, respectively. Next, the total sample was used to compare results between MRI volumes and clinical outcomes when processed via FreeSurfer or FSL. The use, or not, of standardization of volumetric variables was also investigated to assess whether standardizing within sample influenced the effect of the results when comparing different processing software. This was only done in volumetric analysis because the subsequent analysis uses FreeSurfer (for volumes and processing for shape and VBM). Next, the data sites were stratified and meta-analyzed based on MRI field strength and manufacturer to investigate MRI acquisition effects. The total sample was then used to investigate the parallel use of various MRI markers (volume, shape, VBM) across the AD continuum. Lastly, the samples were stratified and meta-analyzed based on ethnicity. All analyses were performed using R (volumetric and shape), Python, and FSLmaths (VBM) (23,24).

Sample Stratification. The samples were stratified, where possible, based on MRI field strength (1.5T/3T), manufacturer (GE/Siemens/Philips), and self-reported ethnicity (African/Asian/European/Hispanic). Regression analyses were run for each stratification using the volumetric data (FreeSurfer), subcortical shape, and VBM data.

Regression Models at Individual Sites. In each study, multivariate logistic (for clinical diagnosis and family history), linear (for analysis assessing *APOE* ε4 load), and Cox (for incident dementia) regressions were performed over 3 models. Model 1 included age and sex; model 2 additionally adjusted for intracranial volume (FreeSurfer variable for all analyses except for FSL models); and model 3 adjusted for age, sex, intracranial volume, and education. Education was adjusted for only in model 3 because these data were not available in the WMH-AD and for 30% of the AIBL participants. In the *APOE* ε4 volumetric analysis, data were analyzed in 2 ways: 1) stratified for diagnosis-free individuals (only including cognitively normal) and 2) additionally adjusted for diagnosis (cognitively normal [0], MCI [1], AD [2]) in the whole sample. The second method was used for the vertex- and voxel-based analysis. To further investigate potential methodological bias, the software analysis was also performed on standardized volumetric variables (*z* matrices). This was only done in the software analysis because all data were utilized together in that analysis. Standardization in the further stratified groups could increase further potential bias.

Meta-Analysis. Results were subsequently meta-analyzed across all samples using the random effects meta-analysis

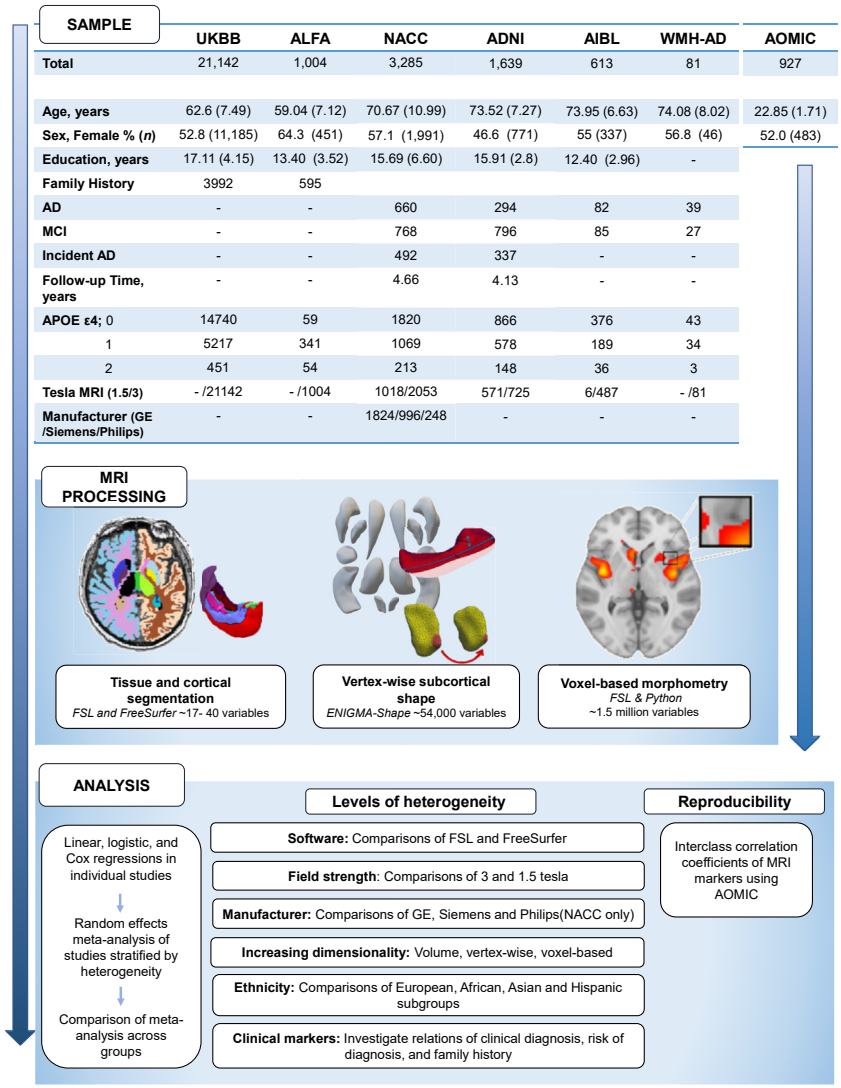

Figure 1. Schematic overview of the study design. AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; AIBL, Australian Imaging Biomarkers and Lifestyle study; ALFA, ALzheimer and FAmilies study; AOMIC, Amsterdam Open MRI Collection; MCI, mild cognitive impairment; MRI, magnetic resonance imaging; NACC, National Alzheimer's Coordinating Center; UKBB, UK Biobank; WMH-AD, White Matter Hyperintensities in Alzheimer's Disease study.

method in the R packages meta and metafor (25). These results were compared between groups using the z test. A positive z score indicates that the effect is larger in this cohort/group than in the reference group, whereas a negative z score indicates that it is smaller. The size of the z score indicates the magnitude of the difference, with a z score of ± 1.64 significant at a .05 p value equivalent. Additional comparison of manufacturers was performed only in the NACC dataset due to large disproportionality across the manufacturers in the data and reasonable homogeneity in the NACC dataset in relation to other factors (15). Similarly, incident AD and family history results are largely restricted to the Patterns Across the AD Continuum section due to little heterogeneity in the data that contains these variables.

Multiple Testing. Bonferroni-adjusted p values were used for volumetric analysis ($p < .0009$) (17 measures, 3 models, .05/51), resulting in a z-value threshold of 3.121. For the shape analysis, the number of independent tests was calculated to be 8591 using UKBB data via permutation testing across all shapes and measures, which resulted in a Bonferroni-adjusted p-value threshold of 5.82×10^{-6} ($z = 4.42$). Similarly, the VBM threshold of $p < 2.99 \times 10^{-7}$ was calculated previously using a permutation method similar to the one that was described previously ($z = 5.06$) (22).

RESULTS

Population Characteristics

The mean ages (in years) of the populations ranged from 62 (UKBB) to 74 (ADNI, AIBL, WMH-AD) across studies (Figure 1). The percentage of the population that was female ranged from 46.6% (ADNI) to 64.3% (ALFA). Mean years of education ranged from 12.40 (AIBL) to 17.11 (UKBB). Notably, the WMH-

AD dataset did not have education information, and in the AIBL, 30% of participants were missing this information. A full overview of the case-control and *APOE* ε4 load distribution is provided in Table 1 and Figure 1 (full results in Table S2 in Supplement 2). For incident AD, the mean (SD) time to follow-up/diagnosis was 4.13 (3.22) and 4.66 (3.32) for the ADNI and NACC, respectively (full results in Table S2 in Supplement 2).

Impact of Data Acquisition and Analysis on Associations

Reproducibility of Imaging in the AOMIC Data. Overall, all methods showed a good level of reproducibility within volumetric segmentation software and the voxel-based measures. However, FreeSurfer had results with substantially better reproducibility than FSL (ICC: 0.89–0.97 compared with 0.58–0.82, respectively). Thus, FreeSurfer was used in all subsequent volumetric analyses. For the subcortical shape analysis, the mean ICC was 0.76 for all 54,000 measures and was particularly high for the thalamus, caudate, putamen, hippocampus, and amygdala. The nucleus accumbens, the smallest subcortical structure, showed the lowest reproducibility (mean ICC: 0.53, percentage above fair: 67.42), while still being considered moderate/fair (23,25). Voxel-based measurements had a mean ICC of 0.95, with 97.9% being above the 0.75 cutoff criterion for "good." For full results, see Table S3 in Supplement 2.

Software (FSL vs. FreeSurfer). When we compared the use of FSL and FreeSurfer, a significant difference in the meta-analyzed results was consistently seen in the relationships between brain volumes and the AD-diagnosis group (vs. the HC group). While a similar pattern of relationships was observed, in model 3, the largest differences in the associations were seen in the amygdala (left, beta FreeSurfer −0.0074; FSL −0.0034) (Figure 2; Table S4 in Supplement 2). For the associations of brain volumes with the MCI-diagnosis group (compared with the HC group), differences in results across the software were seen in volumes of total white matter and the caudate, hippocampus, right thalamus, and putamen (model 3). The largest of these differences was seen in the left hippocampus (z = 11.81). Only the right hippocampus was significantly different across segmentation software in the associations with *APOE* ε4 in the total samples adjusted by diagnosis (Table S4 in Supplement 2).

Standardization. When standardizing the brain volumes within each dataset, the relationships between brain volumes and clinical outcomes across FSL and FreeSurfer had similar patterns of association. Despite this, in the nonstandardized analysis, the left palladium associations did not differ significantly between the 2 processing methods (model 3) (z = 2.59), while in the standardized analysis, they did differ (z = 3.64). Additionally, in the nonstandardized analysis, the left amygdala was significantly different across the software (z = 4.37), while in the standardized analysis it was not different (z = 2.80). For full analysis, see Table S4 in Supplement 2.

Field Strength and Manufacturer. When we examined the data subgrouped by MR field strength and manufacturer, similar patterns of associations were seen across the volumetric relationships with groups with a diagnosis of AD or MCI (each compared with the HC group) across all models. Despite this, differences in the relationships were seen most prominently in the hippocampus (AD: z = 5.39–11.73) and amygdala (AD: z = 5.23–11.61) in model 3 (Figure 2; Tables S5 and S6 in Supplement 2). Differences within these structures were also evident in the shape analysis, especially when we compared GE and Siemens (Tables S5 and S6 in Supplement 2; Figures S2–S5 in Supplement 1). For the field strength comparison, the *APOE* ε4 volumetric analysis in cognitively normal individuals did not yield significantly different relationships; however, when looking at the total sample and adjusting for diagnosis, differences were also seen in relation to the amygdala (left/right z = 3.84/4.67) and the hippocampus (left/right z = 4.40/4.53) (full results in Table S5 in Supplement 2). No differences were observed across MR manufacturers in the

Table 1. Population Subgroups

	APOE ε4 Load			Family History		MCI		AD		Incident AD		
Population Group	0	1	2	No	Yes	No	Yes	No	Yes	No	Yes	*FU Years
Full	18,437	7428	905	14,802	4587	2751	1695	2744	1069	2803	829	4.13–4.66
Ethnicity												
African	235	159	26	80	10	217	80	217	184	213	32	4.11–4.81
Asian	315	100	12	211	28	64	74	64	43	54	16	4.19–5.25
European	17,566	7006	845	14,488	3954	2216	1379	1909	856	2134	722	4.13–5.11
Hispanic	272	126	14	–	–	193	139	193	101	238	46	4.03–5.45
Tesla												
1.5	934	625	131	–	–	669	597	669	418	844	446	4.56–5.77
3	17,271	6577	750	14,802	4587	1986	937	1981	560	1673	370	4.16–4.68
Manufacturer—NACC												
GE	1036	668	141	–	–	1079	419	1076	417	–	–	–
Siemens	557	310	56	–	–	599	248	595	195	–	–	–
Philips	117	53	13	–	–	142	39	142	38	–	–	–

AD, Alzheimer's disease; FU, follow up; MCI, mild cognitive impairment; NACC, National Alzheimer's Coordinating Center.
*FU years display the mean (range) across the samples.

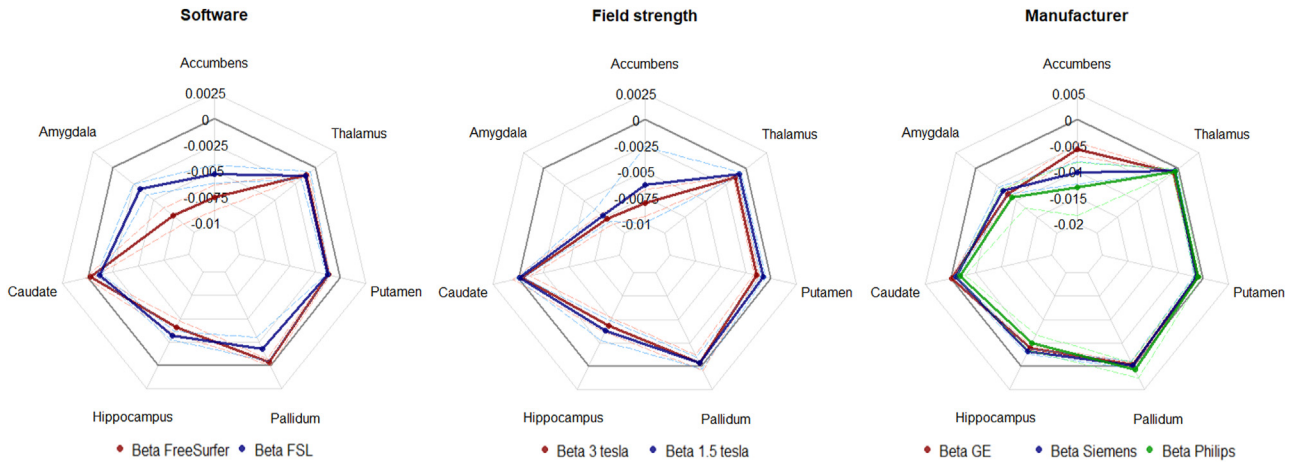

Figure 2. Associations between subcortical brain volumes in the left hemisphere and the Alzheimer's disease group in comparison to the healthy control group across magnetic resonance software and hardware. Bold line represents beta values; dashed lines indicate 95% CIs. Gray line indicates 0. All models are adjusted for age, sex, intracranial volume, and education (model 3). Software analysis: *n* control/case: 2744/1069. Field strength: *n* control/case: 3T, 1981/560; 1.5T, 669/418. Magnetic resonance imaging manufacturer: *n* control/case: GE, 1076/417; Siemens, 595/195, Philips, 142/38.

APOE ε4 analysis. When looking at hippocampal subfield analysis across field strengths and manufacturers, many associations differed significantly across the hardware elements (Table S5 in Supplement 2).

In VBM analysis, similar relationships emerged across MR field strengths and manufacturers across the classifications (Figures S2–S5 in Supplement 1). Clusters of significant differences across field strengths were seen in the thalamus, middle frontal gyrus, parahippocampal area, and insular area in relation to the AD group in comparison to the HC group (Tables S5 and S6 in Supplement 2; Figures S2–S5 in Supplement 1). Similarly, differences in the VBM analysis were mainly observed when GE and Siemens MRI scanners were compared, with the largest voxel clusters of significantly different associations being in the orbital gyrus, substantia nigra, and lateral ventricle frontal horn central part and occipital horn. This latter region also showed differences between Philips and GE/Siemens (Table S6 in Supplement 2).

Relationship Between Imaging Markers and Clinical Outcomes

Insights From Different Imaging Markers. The relationship between imaging markers and the AD group in comparison to the HC group is partly reported in Figure 3 and in full in Table S7 in Supplement 2. In the gross volumetric analysis, all structures were negatively associated with the AD group in comparison with the HC group, most notably the accumbens, amygdala, and hippocampus (larger beta in the left hemisphere). The hippocampal subfield analysis showed that this association was particularly pronounced in the presubiculum, parasubiculum, dentate gyrus, CA4, and CA3 (full results in Table S7 in Supplement 2).

In the shape and VBM analysis, these structures, as well as the anterior and superior temporal gyri, showed the strongest negative associations with the AD group compared with the HC group, indicating volume loss and inward deformation in AD. However, both strong positive and negative clusters were found across the caudate, pallidum, putamen, thalamus, and the insula, which were not identified in the volumetric analyses. For example, in model 3, there was only a weak negative association between AD and right caudate volume (beta [95% CI]: -0.0003 [-6×10^{-4} to -1×10^{-4}]), whereas nearly half of the caudate shape measures (±1191 of 2502) were significantly associated with AD (minimal p value 1.05×10^{-20}), most of which actually showed a positive association (Figure 3; Table S7 in Supplement 2). Similar differential associations, both positive and negative, were seen across other structures and in the MCI and *APOE* ε4 analysis when using different MRI measures (Table S7 in Supplement 2).

Role of Ethnicity. When looking at model 2, due to missing education variables, the largest differences across the non-European ethnic groups in comparison to the European sample were seen when comparing European to African and Hispanic subgroups, specifically for the relationships between AD and the hippocampus (left hippocampus, European-Hispanic $z = 6.54$; European-African $z = 2.95$); this was also observed in the shape analysis (Figure 4; Table S8 in Supplement 2; Figures S6 and S7 in Supplement 1). The comparisons between the European and Hispanic subgroups in model 3 also showed significant differences in the associations with the amygdala, with Europeans showing a stronger negative association (smaller volumes in Europeans with AD) (right amygdala, beta: European -0.0062; Hispanic -0.0045; European-Hispanic $z = 5.28$). The shape analysis showed no significant difference between the relationships across subcortical structure shape and AD diagnosis between the European and Asian groups. The VBM analysis identified small clusters that differed significantly in their relationships to the AD-diagnosis group (compared with the HC group) across European and non-European groups, which were within the lateral ventricle frontal horn central part and occipital horn and the parietal lobe, with larger clusters of voxels being negatively associated with AD in Europeans than other groups (Table S8 in Supplement 2; Figures S6 and S7 in Supplement 1).

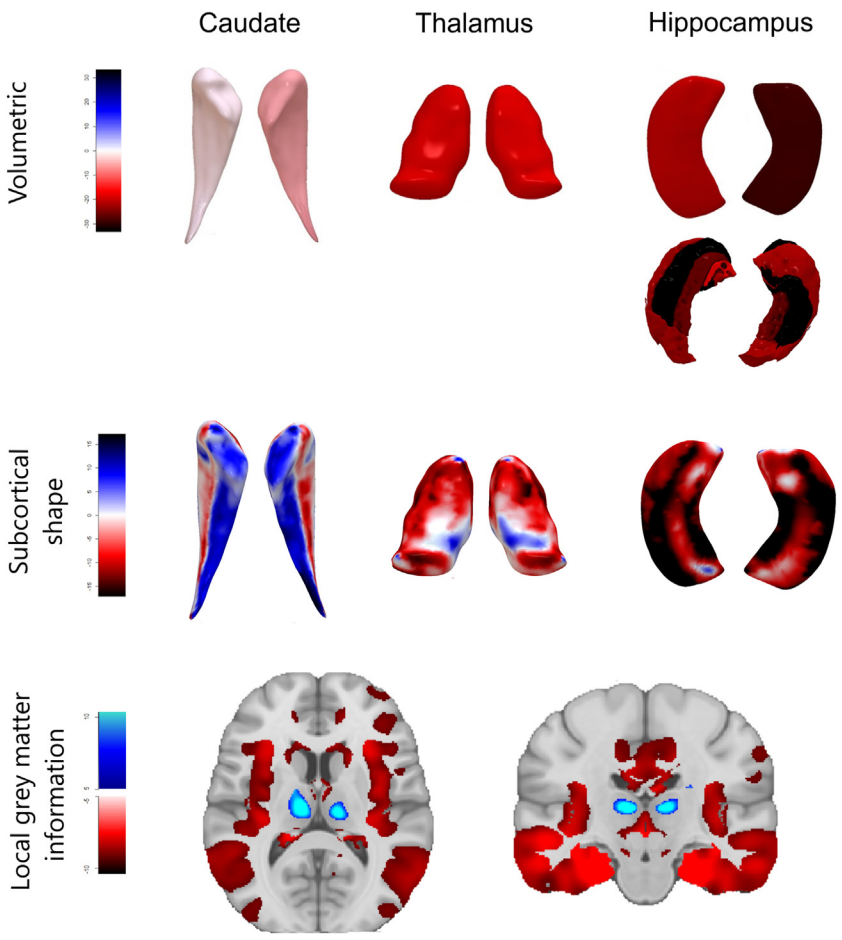

Figure 3. Associations between caudate, thalamus, and hippocampus and the Alzheimer's disease group in comparison to the healthy control group. From top to bottom: volumetric associations including hippocampal subfield volumes, subcortical shape (LogJacs), and voxel-based morphology. All models are adjusted for age, sex, intracranial volume, and education. Color intensity indicates a higher *t* value, with blue indicating larger volume/shape expansion and gray matter intensity and red indicating a smaller volume and shape contraction and lower gray matter intensity. *n* control/case: 2744/1069. Colors represent *t* values with ranges of −32 to 32, −16 to 16, and −10 to 10 from top to bottom, respectively.

The relationship between the MCI-diagnosis group (compared with the HC group) and hippocampal volume was also significantly different between the European and non-European subgroups, predominantly between European and African subgroups (left/right z = 4.69/3.90). No significant differences in the volumetric relations to *APOE* ε4 load were seen across ethnic groups (full results in Table S8 in Supplement 2), except for a small cluster of vertices in the thalamus (European-African comparison).

Patterns Across the AD Continuum. Similar patterns of association were seen, with increasing strength upon increasing clinical severity (Figure 5; full results in Table S7 in Supplement 2). Despite this general increase in relationship strength, the association with AD was most pronounced in the accumbens (beta [95% CI], left −0.0076 [−0.0088 to −0.0064]; right −0.009 [−0.0104 to −0.0076]), amygdala (left −0.0074 [−0.0087 to −0.0062]; right −0.0061 [−0.0071 to −0.0051]) and hippocampus (left −0.0041 [−0.0048 to −0.0034]; right −0.0036 [−0.0039 to −0.0034]), with the relationships with MCI showing similar but more subdued patterns of neurodegeneration. Additionally, MCI diagnosis and incident AD (per year change) had a similar pattern of association and strength. Relationships between the caudate shape were seen in AD and MCI diagnostic groups compared with the HC group; however, this was much smaller or nonexistent when individuals with incident AD were compared with nonconverters (Figure 5; Table S7 in Supplement 2). No relationships were observed for family history of dementia or in the *APOE* ε4 analysis when conducted in the HC group, whereas relationships were observed, although they were small and mainly in the amygdala and hippocampus, when looking in the total sample and adjusting for diagnosis. An example of this increasing strength of the association across the AD disease stages in the left hippocampus are as follows (beta [95% CI]): AD (−0.0041 [−0.0048 to −0.0034]), MCI (−0.0019 [−0.0021 to −0.0017]), incident dementia (per-year increase in risk) (−0.002 [−0.0023 to −0.0018]), *APOE* ε4 (adjusted for diagnosis) (−0.0001693 [−0.0002116 to −0.0001269]), *APOE* ε4 (within cognitively healthy individuals) (−0.0000288 [-4.8×10^{-5} to -9×10^{-6}]), and family history (0.0012 [−0.0012 to 0.0036]).

VBM also showed similar patterns across AD, MCI, and *APOE* ε4, with a surprising larger positive cluster in the thalamus in the AD group than in the HC group (Figure 5; Table S7 in Supplement 2). In the VBM analysis, large negative

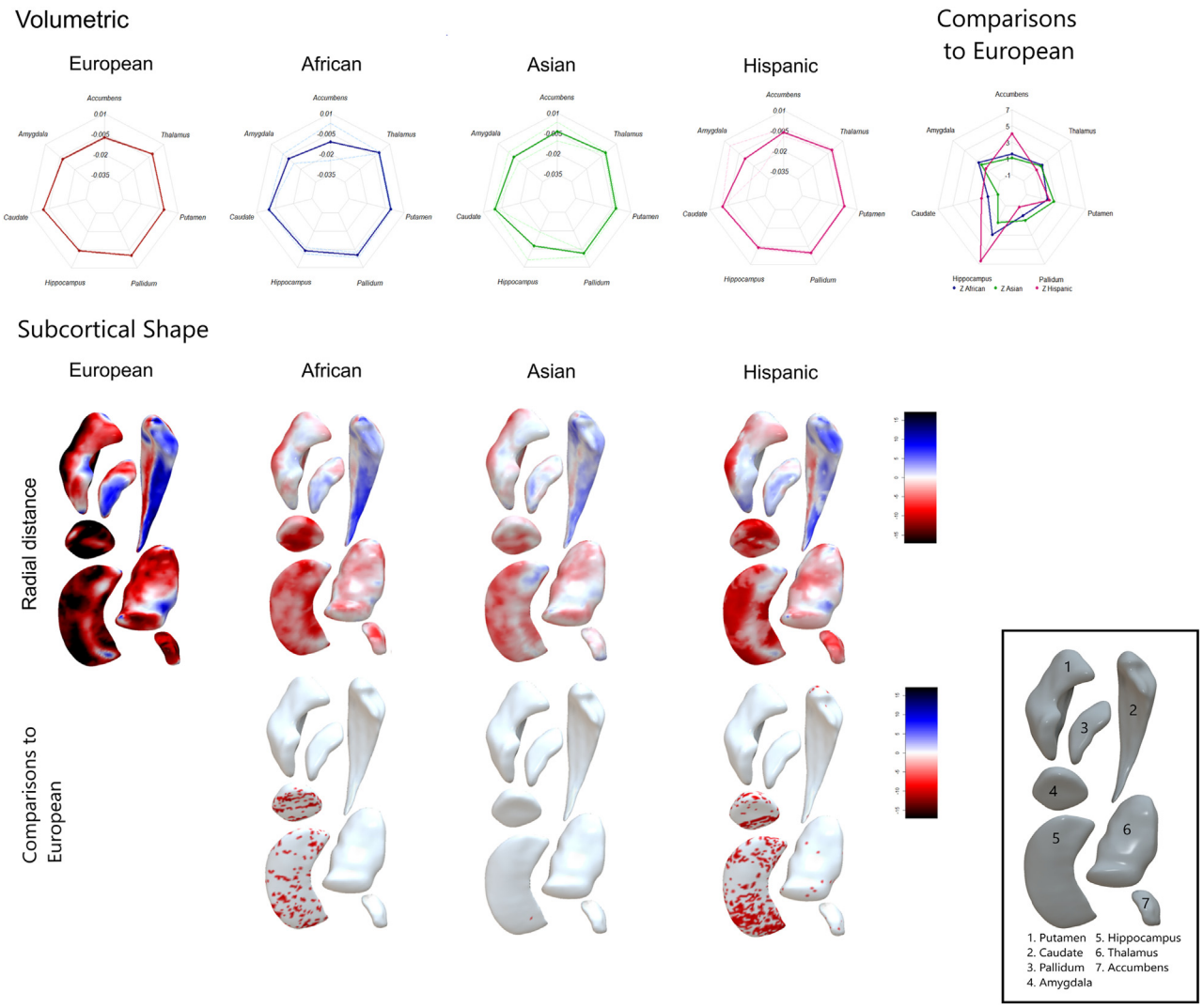

Figure 4. Associations between magnetic resonance markers and the Alzheimer's disease group in comparison to the healthy control group across ethnic groups. Figure depicts left hemisphere associations between regional volumes/subcortical shape with Alzheimer's disease adjusted for age, sex, and intracranial volume. *n* control/case: African, 217/184; Asian, 64/43; European, 1909/856; Hispanic, 193/101. Top: volumetric analysis. Bold line represents beta values, and dashed lines indicate 95% CIs. Bottom: subcortical shape analysis. Radial distance (first row) reports *t* values over the vertices per shape across the ethnicities (range −16 to 16). Color intensity indicates a higher *t* value, with blue indicating larger shape expansion and red indicating a smaller shape contraction than a reference. The second row reports the significant *z* values when comparing each non-European group to the European group (range −16 to 16).

associations with gray matter atrophy and the AD-diagnosis group, compared with the HC group, were observed globally (Table S7 in Supplement 2). Similar associations, although smaller, were observed in the MCI-diagnosis group compared with the HC group. In the current sample, no relationships were seen with family history of dementia across volumetric, shape, and VBM analysis (Table S7 in Supplement 2).

DISCUSSION

In the current study, we investigated the patterns of association between MRI markers and AD stages across different aspects of study design. Many of the association patterns were similar across strata of technical, processing, and clinical groups, which affirms the enormous potential of multisite analysis of heterogeneous data. Despite relatively small non-European groups, the results also revealed subtle differences in patterns of neurodegeneration across ethnic groups.

It is known that differences in data acquisition and processing contribute to variability, and many studies have shown differences in extracted MRI features depending on these factors (6,26–29). In the current study, this also translated to differences in associations with clinical outcomes across AD stages. Despite the overall modest effect, differences in the relationships to the amygdala and hippocampus volumes were observed across technical elements such as field strength,

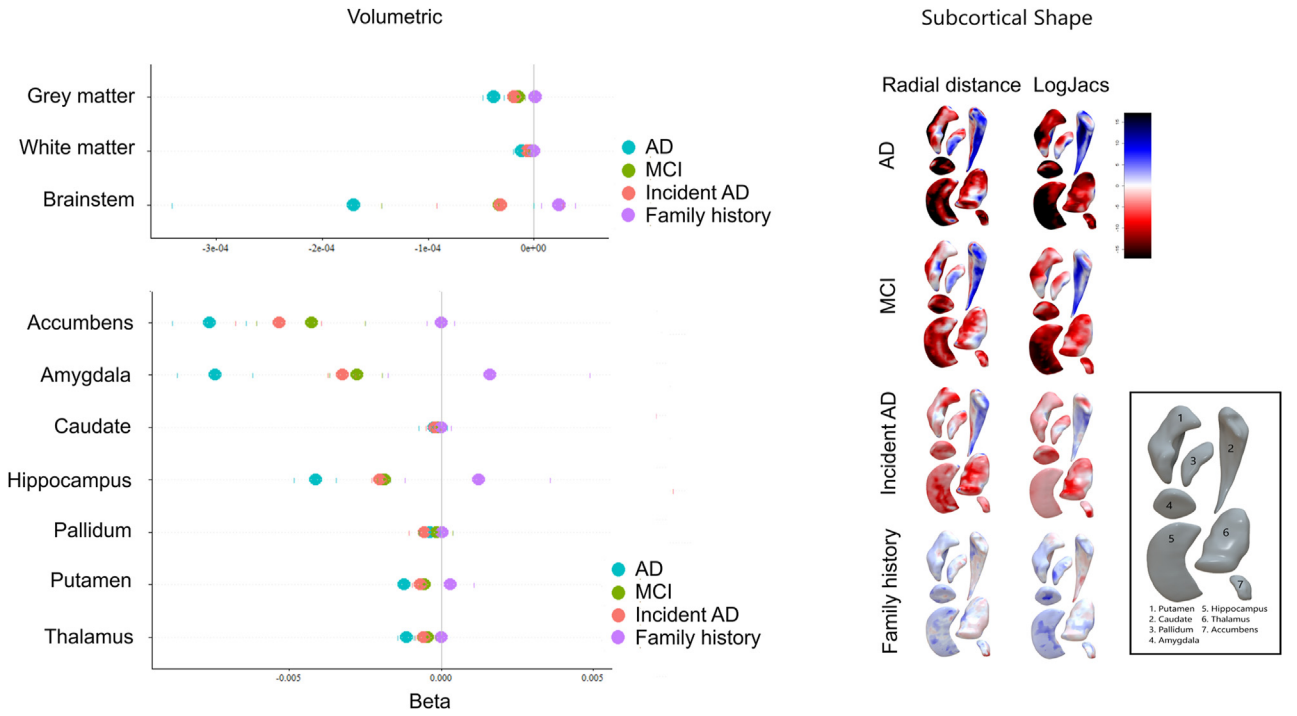

Figure 5. Associations with magnetic resonance markers across clinical variables. Regional volumetric and shape associations with Alzheimer's disease (AD). Volumetric: top, global volumes; bottom, left hemisphere associations. Adjusted for age, sex, intracranial volume, and education. Circles represent beta values, with lines indicating 95% CIs. For incident AD, the circles represent beta as a volume change per year. For the subcortical shape analysis, radial distance and LogJacs are reported from *t* values over the vertices per shape across the classifications. Color intensity indicates a higher *t* value, with blue indicating larger shape expansion and red indicating a shape contraction compared with a reference. *n* control/case: AD, 2744/1069; mild cognitive impairment (MCI), 2751/1695; incident AD, 2803/829; family history, 14,802/4587.

despite other studies finding little effects (26). When segmenting small regions, such as the hippocampal subregions, differences in relationships to AD became more apparent across MR hardware. Thus, care must be taken when investigating small regions within heterogeneous MR sources, and other methods, such as voxel-based analysis, may be preferable. Other large differences were seen when software types were compared, namely FreeSurfer and FSL segmentation pipelines. Thus, efforts toward harmonization of processing software are crucial when meta-analyzing multisite data, especially with the field moving toward more fine-grained metrics. Harmonized processing in fine-grained metrics also enables the gold-standard approach of meta-analyzing at each vertex and voxel, whereas the less ideal peak-based meta-analysis techniques are often used when meta-analyzing published work and/or differing methods (30). Despite this, caution should be taken regarding multisite VBM due to potential gray matter volume differences; however, meta-analysis at the summary statistic level may overcome these potential problems (31). Additionally, there are preanalysis harmonization tools that aim to estimate and correct for heterogeneity (32). However, while promising, these tools have their own limitations, including lack of reliability in small samples and difficulty being applied to multisite data that are not shared and/or without access to high-performance computing infrastructure (33). Furthermore, recent research has shown that some methods can actually harm the analysis when using structural imaging data, with one study observing that harmonization methods, such as neuroCombat, appeared to obscure some true biological effects when applied to structural MRI (33). Thus, while extremely useful for many situations, in some instances, such methods need further development. Altogether, these results indicate that postacquisition harmonization methods cannot completely negate differences in image collection across all situations.

While measurement heterogeneity can result in some potential bias, we do not conclude that uniform data is the solution. Global harmonization of neuroimaging machines and protocols in clinical practice is not realistic with new machines and protocols being developed constantly and with disparities in access to these technologies in low- and middle-income countries. Thus, the use of heterogeneous data in research results in more generalizable conclusions with better real-word translation. Furthermore, there are sometimes specific choices for machines that could not be addressed with uniformity, for example, a weaker magnet ($<$1.5T) when a patient has foreign bodies in their brain (34) or the use of low-field MRI or computed tomography (CT) when standard MRI is not feasible. The modest differences that we observed when comparing 1.5T and 3T MRI scanners are likely to increase when including more heterogeneous imaging modalities such as CT or low-field MRI. While this may be true, current consortia have combined CT and MR information and arrived at the same conclusions when studying relatively crude measures such as

intracranial volume (35). Despite similar results being reported in MR investigations when using CT-segmented gray matter and white matter in relation to markers of neurodegeneration, combining different modalities such as CT and MRI with more fine-grained variables needs further testing (36). Although it is unknown what the impact of even more heterogeneous acquisition will be, these points, together with the minimal differences in patterns seen in the current study inclusive of heterogeneous data, suggest that it is time to accept that no singular way of acquiring data is optimal and that the use of varied data will have a wider impact.

While homogeneity of methods that extract specific measurements is preferred, having heterogeneity in processing and analysis can also be valuable. The use of multiple processing methods to gather different, but complementary, information provides more insight into the atrophy patterns of diseases (37). The current investigation into the relationship between caudate volume and AD yielded a modest negative association. When investigating this region with more detailed vertex- and voxel-based measures, strong negative and positive associations to AD were both observed. This suggests that regions in the caudate have differential associations to the disease, which is missed with gross volumetric analysis. The caudate nucleus itself is a complex structure, with different portions of the caudate receiving projections from distinct cortical regions and therefore also being involved in different cognitive processes. Thus, these more detailed imaging measures allow a more complete depiction of the neurodegenerative process and its progression in the brain.

Despite relatively small non-European samples, the current study found subtle variations across ethnic groups, with differential patterns emerging across detailed imaging markers. Few previous studies have investigated differences in the relationship between brain atrophy, AD, and cognitive decline across populations, and those studies have reported mixed results, potentially due to cohort and methodological differences (38–42). In the volumetric analysis, negative associations between AD and MCI and the hippocampus and accumbens were significantly different in non-Europeans and Europeans. There were no significant differences in the strength of associations in the Asian stratum compared with the other groups, but this is not conclusive due to the small sample size and large variation within this stratum. Additionally, due to education information missing from the Taiwan site, we mostly studied ethnic differences in model 2 that were not adjusted for education. Education and other unmeasured confounders could potentially further influence these relationships. This initial investigation is the first step for larger studies planned in the wider UNITED consortium. The current results provide support for the use of detailed MRI markers to unravel subtle similarities and differences in disease progress across populations.

Notably, almost no significant differences were found in the effects of the *APOE* ε4 alleles in the ethnic groups. This lack of differentiation is interesting given the reports that despite higher rates of prevalence of the ε4 allele in Africans, less effect of the ε4 allele on risk for AD is observed. However, the largest effect of *APOE* ε4 on factors such as AD risk occurs in Europeans (43). Relatively few studies have investigated the direct effect of this ε4 allele on brain architecture across ethnic groups; however, one study suggested a differential effect of *APOE* ε4 on brain volume in Chinese versus White individuals (44). Conflicting results have been found with positron emission tomography. Despite both higher positron emission tomography positivity and higher standardized uptake value ratios being related to the ε4 allele, research has found that African Americans have lower rates of positron emission tomography positivity than Whites but also higher odds of elevated standardized uptake value ratios (45,46). We believe that larger, multisite efforts will help resolve conflicting findings, with limited, single-site efforts typically being underpowered. Nevertheless, reports of differences in *APOE* alleles between populations could also be explained by other genetic or environmental factors that modify effects. In future research, it is important to measure these potential modifiers to understand the pathways in which these effects occur. For example, the current sites were largely based in European, North American, and Australian locations, and it is unclear to what extent these findings can be extrapolated to other settings with their unique local backdrops.

The current investigation highlights the potential of incorporating data from multiple sites and heterogeneous origins, including both the ability to replicate previously found associations in single sites despite heterogeneous data and the ability to tease apart subtle differences across populations. Despite only using a sample of the full consortium, strengths of this study include its size and the systematic nature of the investigation to uncover potential bias and its impact within this study design. Furthermore, the ability to use the raw output for the high-dimensional markers for the meta-analysis is beneficial and will allow for more accuracy in collaborative full consortium work. This also allowed for direct comparisons of groups at the voxel and vertex levels, which considers multiple statistical elements rather than only the significant cluster size and location as is done with peak-based meta-analysis.

Our work also outlines what is needed in the future in large multisite consortia. Importantly, the inclusion of more diverse populations (i.e., non-European), which is the main focus of UNITED, is lacking in the field (47–49). Related to this, our current analysis focused on specific acquisition protocols that are typical in Europe and North America, but more variability exists. Other elements, such as multimodal acquisition with CT and low-field MRI may be more difficult to harmonize and conduct joint analyses on. In addition, more variability exists in image processing that was not assessed in the current project, because our focus was on methods that are open source and therefore more accessible. However, the results do depict the benefits of harmonizing software when extracting the same markers. On the clinical side, comparisons may be complicated by diagnostic approaches, including mixed pathology with single classifications due to a lack of biomarker-defined diagnosis. Other limitations of the current analysis include potential further confounding across strata whereby the observed differences may be influenced by factors other than those under investigation, such as ethnicity in the broader investigations.

Conclusions

The current study shows great potential for the wider consortium and other international multisite collaborations. The

investigations highlight the importance of implementing harmonized processing and analysis pipelines to alleviate potential bias but also the importance of being transparent and stratifying more on aspects of study design when harmonization is not an option. This will not only allow for more accurate interpretation of the results but will also allow for better generalization across clinical populations. Further research in the consortium will allow for robust results and subgroup investigations while retaining high levels of statistical power, especially in high-dimensional imaging analysis, which requires strict thresholding. Furthering this work has the potential to investigate imaging of neurodegeneration across a wide range of diverse population groups together.

ACKNOWLEDGMENTS AND DISCLOSURES

TEE is funded by the Alzheimer's Nederland WE.03-2020-08, NWO (Grant No. 19501). NV-T is funded by the Juan de la Cierva Programme (Grant No. FJC2018-038085-I), Ministerio de Ciência, Innovación y Universidades–Spanish State Research Agency (IJC2020-043216-I/MCIN/AEI/10.13039/501100011033), the European Union Next Generation EU/PRTR and currently recieves funding from grant RYC2022-038136-I cofunded by the European Union FSE+ and grant PID2022-143106OA-100 cofunded by the European Union FEDER. Additionally NV-T is supported by the William H Gates Sr Fellowship from the Alzheimer's Disease Data Initative, La Caxia Foundation and the Barcelona City Council (project code: 23S06083-001). 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. AI is supported by grants from CONICET; ANID/FONDECYT Regular (Grant Nos. 1210195, 1210176, and 1220995); ANID/FONDAP/15150012; ANID/PIA/ANILLOS ACT210096; FONDEF ID20I10152, ID22I10029; ANID/FONDAP 15150012; Takeda CW2680521 and the MULTI-PARTNER CONSORTIUM TO EXPAND DEMENTIA RESEARCH IN LATIN AMERICA [ReDLat, supported by Fogarty International Center, National Institutes of Health (NIH), National Institutes of Aging (Grant Nos. R01 AG057234, R01 AG075775, R01 AG21051, CARDS-NIH), Alzheimer's Association (Grant No. SG-20-725707), Rainwater Charitable Foundation–The Bluefield project to cure FTD, and Global Brain Health Institute]. JDG is supported by the Spanish Ministry of Economy and Competitiveness (Grant No. RYC-2013-13054) and received research support from the EU/EFPIA Innovative Medicines Initiative Joint Undertaking AMYPAD (Grant No. 115952) and Ministerio de Ciencia, Innovación y Universidades (Grant No. RTI2018-102261). HHHA was supported by ZonMw (Grant No. 916.19.151).

ALFA: The research that led to these results received funding from la Caixa Foundation (ID 100010434), under agreement LCF/PR/GN17/10300004 and the Alzheimer's Association, and an international anonymous charity foundation through the TriBEKa Imaging Platform project (Grant No. TriBEKa-17-519007). Additional support has been recieved from the Universities and Research Secretariat, Ministry of Business and Knowledge of the Catalan Government under Grant no. 2021 SGR 00913.

UKBB: The UKBB has ethics approval from the North West Multi-Centre Research Ethics Committee and possesses a Human Tissue Authority license. This study was carried out using UKBB Application number 23509, and we thank the participants in the UKBB imaging study (https://www.ukbiobank.ac.uk/).

ADNI: Data collection and sharing for this project was funded by the ADNI (NIH Grant No. U01 AG024904) and DOD ADNI (Department of Defense award No. 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 NIH (http://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.

AIBL: Funding for the study was provided in part by the study partners (Commonwealth Scientific Industrial and Research Organization, Edith Cowan University, Mental Health Research Institute, National Ageing Research Institute, Austin Health, CogState Ltd.). The study also received support from the National Health and Medical Research Council and the Dementia Collaborative Research Centres program, as well as funding from the Science and Industry Endowment Fund and the Cooperative Research Centre for Mental Health, an Australian Government Initiative.

NACC: The NACC database is funded by NIA/NIH (Grant No. U24 AG072122). NACC data are contributed by the NIA-funded ADCs: P50 AG005131 (principal investigator [PI] James Brewer, M.D., Ph.D.), P50 AG005133 (PI Oscar Lopez, M.D.), P50 AG005134 (PI Bradley Hyman, M.D., Ph.D.), P50 AG005136 (PI Thomas Grabowski, M.D.), P50 AG005138 (PI Mary Sano, Ph.D.), P50 AG005142 (PI Helena Chui, M.D.), P50 AG005146 (PI Marilyn Albert, Ph.D.), P50 AG005681 (PI John Morris, M.D.), P30 AG008017 (PI Jeffrey Kaye, M.D.), P30 AG008051 (PI Thomas Wisniewski, M.D.), P50 AG008702 (PI Scott Small, M.D.), P30 AG010124 (PI John Trojanowski, M.D., Ph.D.), P30 AG010129 (PI Charles DeCarli, M.D.), P30 AG010133 (PI Andrew Saykin, Psy.D.), P30 AG010161 (PI David Bennett, M.D.), P30 AG012300 (PI Roger Rosenberg, M.D.), P30 AG013846 (PI Neil Kowall, M.D.), P30 AG013854 (PI Robert Vassar, Ph.D.), P50 AG016573 (PI Frank LaFerla, Ph.D.), P50 AG016574 (PI Ronald Petersen, M.D., Ph.D.), P30 AG019610 (PI Eric Reiman, M.D.), P50 AG023501 (PI Bruce Miller, M.D.), P50 AG025688 (PI Allan Levey, M.D., Ph.D.), P30 AG028383 (PI Linda Van Eldik, Ph.D.), P50 AG033514 (PI Sanjay Asthana, M.D., F.R.C.P.), P30 AG035982 (PI Russell Swerdlow, M.D.), P50 AG047266 (PI Todd Golde, M.D., Ph.D.), P50 AG047270 (PI Stephen Strittmatter, M.D., Ph.D.), P50 AG047366 (PI Victor Henderson, M.D., M.S.), P30 AG049638 (PI Suzanne Craft, Ph.D.), P30 AG053760 (PI Henry Paulson, M.D., Ph.D.), P30 AG066546 (PI Sudha Seshadri, M.D.), P20 AG068024 (PI Erik Roberson, M.D., Ph.D.), P20 AG068053 (PI Marwan Sabbagh, M.D.), P20 AG068077 (PI Gary Rosenberg, M.D.), P20 AG068082 (PI Angela Jefferson, Ph.D.), P30 AG072958 (PI Heather Whitson, M.D.), P30 AG072959 (PI James Leverenz, M.D.).

WMH-AD: This work was supported by the Global Alzheimer's Association Interactive Network initiative of the Alzheimer's Association (Grant No. GAAIN-14-244631) and NIH (Grants Nos. U54-EB020406 and P41-EB015922). WMH-AD data used in the preparation of this article were obtained from neuGRID platform (https://neugrid2.eu) hosted in the HPC center of the IRCCS Centro San Giovanni di Dio Fatebenefratelli funded by the European (FP7/2007-2013 Grant agreement No. 283562) and National initiatives (RC-2016-2361095, RRC-2017-2364915, RRC-2018-2365796, CCR-2017-23669078, RRC-2019-23669119_001). The contents of this publication are solely the responsibility of the authors and do not represent the official views of these institutions. All human subjects provided informed consent within each individual study.

TEE, AH, AI, SS, LCST, MW, SA, UA, and HHHA were responsible for concept conception; TEE and HHHA conceived and designed the analysis; TEE was responsible for data processing; TEE performed the analysis; TEE and HHHA wrote the manuscript; NV-T, GO, CF, and JDG contributed data or analysis tools; all authors were responsible for manuscript revisions.

We thank SURFsara (http://www.surfsara.nl) for their support in using the Dutch National Supercomputer Cartesius and through NWO Grant No. 2019.014.

Data used in preparation of this article were obtained from the 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. Data used in the preparation of this article was obtained from the AIBL funded by the Commonwealth Scientific and Industrial Research Organization which was made available at the ADNI database (http://www.loni.usc.edu/ADNI). The AIBL researchers contributed data but did not participate in analysis or writing of this report. AIBL researchers are listed at http://www.aibl.csiro.au. All data except for ALFA is open source and can be applied for through their respective sites. For ALFA information regarding submitting proposals and accessing data may be requested at http://www.barcelonabeta.org/en/contact.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Clinical Genetics, Erasmus MC, Rotterdam, the Netherlands (TEE, NV-T, HHHA); Department of Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, the Netherlands (TEE, HHHA); BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation, Barcelona, Spain (NV-T, GO, CF, JDG); Centre for Genomic Regulation, The Barcelona Institute for Science and Technology, Barcelona, Spain (NV-T); IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain (NV-T); Centro de Investigación Biomédica en Red Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), Madrid, Spain (CF); Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (AH); Latin American Brain Health Institute (BrainLat), Universidad Adolfo Ibáñez, Santiago de Chile, Santiago, Peñalolén, Región Metropolitana, Chile (AI, HHHA); Universidad de San Andrés & Consejo Nacional de Investigaciones Científicas y técnicas, Victoria, Provincia de Buenos Aires, Argentina (AI); Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland (AI); Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, University of Texas Health Science Center, San Antonio, Texas (SS); Department of Neurology, National Neuroscience Institute, Singapore, Singapore (LCST); Parkinson's Disease and Movement Disorders Centre, International Centre of Excellence, USA Parkinson Foundation, Singapore, Singapore (LCST); Department of Veterans Affairs Medical Center, Center for Imaging of Neurodegenerative Diseases, VA Medical Center, San Francisco, California (MW); Department of Neurology, University of California, San Francisco, California (MW); Department of Neurology, National Institute of Mental Health and Neurosciences, Bengaluru, India (SA); and Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada (UA).

Address correspondence to Hieab H.H. Adams, M.D., Ph.D., at h.adams@erasmusmc.nl.

Received Apr 26, 2024; revised Jun 8, 2024; accepted Jul 9, 2024.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2024.07.019.

REFERENCES

1. GBD 2019 Dementia Forecasting Collaborators (2022): Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the Global Burden of Disease Study 2019. Lancet Public Health 7:e105–e125.
2. Prince MJ, Wimo A, Guerchet MM, Ali GC, Wu Y-T, Prina M (2015): World Alzheimer Report 2015-The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. Available at: https://www.alzint.org/resource/world-alzheimer-report-2015/. Accessed July 1, 2022.
3. Gorgolewski KJ, Poldrack RA (2016): A practical guide for improving transparency and reproducibility in neuroimaging research. PLoS Biol 14:e1002506.
4. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR (2013): Power failure: Why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14:365–376.
5. Open Science Collaboration (2015): PSYCHOLOGY. Estimating the reproducibility of psychological science. Science 349:aac4716.
6. Greene AS, Shen X, Noble S, Horien C, Hahn CA, Arora J, *et al.* (2022): Brain–phenotype models fail for individuals who defy sample stereotypes. Nature 609:109–118.
7. Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, *et al.* (2010): Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. Neurology 74:201–209.
8. Parra MA, Baez S, Sedeño L, Gonzalez Campo C, Santamaría-García H, Aprahamian I, *et al.* (2021): Dementia in Latin America: Paving the way toward a regional action plan. Alzheimers Dement 17:295–313.
9. Adams HHH, Evans TE, Terzikhan N (2019): The uncovering neurodegenerative insights through ethnic diversity consortium. Lancet Neurol 18:915.
10. Thompson PM, Jahanshad N, Ching CR, Salminen LE, Thomopoulos SI, Bright J, *et al.* (2020): ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. Transl Psychiatry 10:100.
11. Sargurupremraj M, Suzuki H, Jian X, Sarnowski C, Evans TE, Bis JC, *et al.* (2020): Cerebral small vessel disease genomics and its implications across the lifespan. Nat Commun 11:6285.
12. Knol MJ, Lu D, Traylor M, Adams HHH, Romero JRJ, Smith AV, *et al.* (2020): Association of common genetic variants with brain microbleeds: A genome-wide association study. Neurology 95:e3331–e3343.
13. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, *et al.* (2015): UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 12:e1001779.
14. Molinuevo JL, Gramunt N, Gispert JD, Fauria K, Esteller M, Minguillon C, *et al.* (2016): The ALFA project: A research platform to identify early pathophysiological features of Alzheimer's disease. Alzheimers Dement (N Y) 2:82–92.
15. Besser L, Kukull W, Knopman DS, Chui H, Galasko D, Weintraub S, *et al.* (2018): Version 3: of the National Alzheimer's Coordinating Center's Uniform data set. Alzheimer Dis Assoc Disord.
16. Ellis KA, Bush AI, Darby D, De Fazio D, Foster J, Hudson P, *et al.* (2009): The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. Int Psychogeriatr 21:672–687.
17. Tu M-C, Huang C-W, Chen N-C, Chang W-N, Lui C-C, Chen C-F, *et al.* (2010): Hyperhomocysteinemia in Alzheimer dementia patients and cognitive decline after 6 months follow-up period. Acta Neurol Taiwan 19:168–177.
18. Snoek L, van der Miesen MM, Beemsterboer T, van der Leij A, Eigenhuis A, Scholte HS (2021): The Amsterdam Open MRI Collection, a set of multimodal MRI datasets for individual difference analyses. Sci Data 8:85.
19. Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, *et al.* (2015): A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. Neuroimage 115:117–137.
20. Patenaude B, Smith SM, Kennedy DN, Jenkinson M (2011): A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage 56:907–922.
21. Gutman BA, Wang Y, Rajagopalan P, Toga AW, Thompson PM, editors (2012): Shape matching with medial curves and 1-D group-wise registration. In: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI). Barcelona, Spain: IEEE, 716–719.
22. Roshchupkin GV, Adams HH, van der Lee SJ, Vernooij MW, van Duijn CM, Uitterlinden AG, *et al.* (2016): Fine-mapping the effects of Alzheimer's disease risk loci on brain morphology. Neurobiol Aging 48:204–211.
23. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, *et al.* (2004): Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23(suppl 1):S208–S219.
24. Van Rossum G, Drake FL (2000): Python Reference Manual: iUniverse Indiana. Available at: https://lab.demog.berkeley.edu/Docs/Refs/Python2.1/ref.pdf. Accessed February 1, 2020.

25. Balduzzi S, Rücker G, Schwarzer G (2019): How to perform a meta-analysis with R: A practical tutorial. Evid Based Ment Health 22: 153–160.
26. Heinen R, Bouvy WH, Mendrik AM, Viergever MA, Biessels GJ, De Bresser J (2016): Robustness of automated methods for brain volume measurements across different MRI field strengths. PLoS One 11: e0165719.
27. Haddad E, Pizzagalli F, Zhu AH, Bhatt RR, Islam T, Ba Gari I, *et al.* (2023): Multisite test–retest reliability and compatibility of brain metrics derived from FreeSurfer versions 7.1, 6.0, and 5.3. Hum Brain Mapp 44:1515–1532.
28. Srinivasan D, Erus G, Doshi J, Wolk DA, Shou H, Habes M, *et al.* (2020): A comparison of FreeSurfer and multi-atlas MUSE for brain anatomy segmentation: Findings about size and age bias, and inter-scanner stability in multi-site aging studies. Neuroimage 223: 117248.
29. Cover KS, van Schijndel RA, Versteeg A, Leung KK, Mulder ER, Jong RA, *et al.* (2016): Reproducibility of hippocampal atrophy rates measured with manual, FreeSurfer, AdaBoost, FSL/FIRST and the MAPS-HBSI methods in Alzheimer's disease. Psychiatry Res Neuroimaging 252:26–35.
30. Müller VI, Cieslik EC, Laird AR, Fox PT, Radua J, Mataix-Cols D, *et al.* (2018): Ten simple rules for neuroimaging meta-analysis. Neurosci Biobehav Rev 84:151–161.
31. Focke NK, Helms G, Kaspar S, Diederich C, Tóth V, Dechent P, *et al.* (2011): Multi-site voxel-based morphometry—Not quite there yet. Neuroimage 56:1164–1170.
32. Fortin J-P, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, *et al.* (2018): Harmonization of cortical thickness measurements across scanners and sites. Neuroimage 167:104–120.
33. Richter S, Winzeck S, Correia MM, Kornaropoulos EN, Manktelow A, Outtrim J, *et al.* (2022): Validation of cross-sectional and longitudinal ComBat harmonization methods for magnetic resonance imaging data on a travelling subject cohort. Neuroimage Rep 2:None.
34. Wood R, Bassett K, Foerster V, Spry C, Tong L (2012): 1.5 tesla magnetic resonance imaging scanners compared with 3.0 tesla magnetic resonance imaging scanners: Systematic review of clinical effectiveness. CADTH Technol Overv 2:e2201.
35. Knol MJ, Poot RA, Evans TE, Satizabal CL, Mishra A, Sargurupremraj M, *et al.* (2024): Genetic variants for head size share genes and pathways with cancer. Cell Rep Med 5:10152.
36. Srikrishna M, Ashton NJ, Moscoso A, Pereira JB, Heckemann RA, Van Westen D, *et al.* (2024): CT-based volumetric measures obtained through deep learning: Association with biomarkers of neurodegeneration. Alzheimers Dement 20:629–640.
37. Bento M, Fantini I, Park J, Rittner L, Frayne R (2021): Deep learning in large and multi-site structural brain MR imaging datasets. Front Neuroinform 15:805669.
38. Fan J, Tse M, Carr JS, Miller BL, Kramer JH, Rosen HJ, *et al.* (2019): Alzheimer disease-associated cortical atrophy does not differ between Chinese and whites. Alzheimer Dis Assoc Disord 33:186–193.
39. Zahodne LB, Manly JJ, Narkhede A, Griffith EY, DeCarli C, Schupf NS, *et al.* (2015): Structural MRI predictors of late-life cognition differ across African Americans, Hispanics, and Whites. Curr Alzheimer Res 12:632–639.
40. Gavett BE, Fletcher E, Harvey D, Farias ST, Olichney J, Beckett L, *et al.* (2018): Ethnoracial differences in brain structure change and cognitive change. Neuropsychology 32:529–540.
41. Burke SL, Rodriguez MJ, Barker W, Greig-Custo MT, Rosselli M, Loewenstein DA, Duara R (2018): Relationship between cognitive performance and measures of neurodegeneration among Hispanic and white non-Hispanic individuals with normal cognition, mild cognitive impairment, and dementia. J Int Neuropsychol Soc 24:176–187.
42. Gu Y, Razlighi QR, Zahodne LB, Janicki SC, Ichise M, Manly JJ, *et al.* (2015): Brain amyloid deposition and longitudinal cognitive decline in nondemented older subjects: Results from a multi-ethnic population. PLoS One 10:e0123743.
43. Qin W, Li W, Wang Q, Gong M, Li T, Shi Y, *et al.* (2021): Race-related association between APOE genotype and Alzheimer's disease: A systematic review and meta-analysis. J Alzheimers Dis 83:897–906.
44. Yokoyama JS, Lee AKL, Takada LT, Busovaca E, Bonham LW, Chao SZ, *et al.* (2015): Apolipoprotein ε4 is associated with lower brain volume in cognitively normal Chinese but not white older adults. PLoS One 10:e0118338.
45. Wilkins CH, Windon CC, Dilworth-Anderson P, Romanoff J, Gatsonis C, Hanna L, *et al.* (2022): Racial and ethnic differences in amyloid PET positivity in individuals with mild cognitive impairment or dementia: A secondary analysis of the imaging dementia–evidence for amyloid scanning (IDEAS) cohort study. JAMA Neurol 79:1139–1147.
46. Gottesman RF, Schneider ALC, Zhou Y, Chen X, Green E, Gupta N, *et al.* (2016): The ARIC-PET amyloid imaging study: Brain amyloid differences by age, race, sex, and APOE. Neurology 87:473–480.
47. (2022): 2022 Alzheimer's disease facts and figures. Alzheimers Dem 18:700–789.
48. Brett BL, Schneider JA, Aggarwal NT (2021): Diversity in aging-related neuroimaging research. Pract Neurol 44–46.
49. Babulal GM, Quiroz YT, Albensi BC, Arenaza-Urquijo E, Astell AJ, Babiloni C, *et al.* (2019): Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: Update and areas of immediate need. Alzheimers Dement 15:292–312.