HHS Public Access

Author manuscript

Neuroimage Rep. Author manuscript; available in PMC 2025 January 30.

Published in final edited form as:

Neuroimage Rep. 2024 December ; 4(4): . doi:10.1016/j.ynirp.2024.100227.

MRI-guided clustering of patients with mild dementia due to Alzheimer's disease using self-organizing maps

Kellen K. Petersen[a,*], **Bhargav T. Nallapu**[b], **Richard B. Lipton**[b], **Ellen Grober**[b], **Ali Ezzati**[b,c]

[a]Department of Neurology, Washington University in St. Louis, St. Louis, MO, USA

[b]The Saul R. Korey Department of Neurology, Albert Einstein College of Medicine, New York City, NY, USA

[c]Department of Neurology, University of California-Irvine, Irvine, CA, USA

Abstract

Introduction: Alzheimer's disease (AD) is a phenotypically and pathologically heterogenous neurodegenerative disorder. This heterogeneity can be studied and disentangled using data-driven clustering techniques.

Methods: We implemented a self-organizing map clustering algorithm on baseline volumetric MRI measures from nine brain regions of interest (ROIs) to cluster 1041 individuals enrolled in the placebo arm of the EXPEDITION3 trial. Volumetric MRI differences were compared among clusters. Demographics as well as baseline and longitudinal cognitive performance metrics were used to evaluate cluster characteristics.

Results: Three distinct clusters, with an overall silhouette coefficient of 0.491, were identified based on MRI volumetrics. Cluster 1 (N = 400) had the largest baseline volumetric measures across all ROIs and the best cognitive performance at baseline. Cluster 2 (N = 269) had larger hippocampal and medial temporal lobe volumes, but smaller parietal lobe volumes in comparison with the third cluster (N = 372). Significant between-group mean differences were observed between Clusters 1 and 2 (difference, 2.38; 95% CI, 1.85 to 2.91; P < 0.001), Clusters 1 and 3 (difference, 1.93; 95% CI, 1.41 to 2.44; P < 0.001), but not between Clusters 2 and 3 (difference, 0.45; 95% CI, −0.11 to 1.02; P = 0.146) in ADAS-14.

*Corresponding author. kellenpetersen@Wustl.edu (K.K. Petersen).

CRediT authorship contribution statement

Kellen K. Petersen: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Bhargav T. Nallapu:** Writing – review & editing, Investigation, Conceptualization. **Richard B. Lipton:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Ellen Grober:** Writing – review & editing, Supervision, Funding acquisition. **Ali Ezzati:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ynirp.2024.100227.

Conclusions: Volumetric MRI can be used to identify homogenous clusters of amyloid positive individuals with mild dementia. The groups identified differ in baseline and longitudinal characteristics. Cluster 1 shows little ADAS-14 change over the first 40 weeks of study on placebo treatment and may be unsuitable for identifying early benefits of treatment.

Keywords

Alzheimer's disease; Structural MRI; Unsupervised learning; Machine learning; Subtypes; Mild dementia

1. INTRODUCTION

Alzheimer's disease (AD) is a neurodegenerative condition that exhibits phenotypical and pathological heterogeneity. The recent introduction of ATN research framework (Jack et al., 2018), a clinically agnostic conceptual framework focused on a biological definition of AD is expected to change the landscape of clinical trials through pathway-based therapeutic approaches. The current framework is largely based on using biomarkers cutoffs for stratification of individuals. It is expected that modifications to the ATN framework will be made through the addition of other pathophysiological process, including neuroinflammation, vascular damage, and accounting for spatial distribution of pathology (Gauthier et al., 2018).

Hypothesis-driven studies have successfully identified atrophy-related subtypes. However, this approach is limited due to its reliance on prior knowledge and definitions of neuropathological subtypes (Habes et al., 2020). Data-driven methods combined with advances in neuroimaging, on the other hand, offer the opportunity to find novel subtypes from in-vivo spatial patterns of pathology within diverse patient populations (Habes et al., 2020). Individuals with similar spatial patterns of AD pathology are more likely to show the same clinical course (e.g., cognitive decline) and are more likely to response to the same treatment (Ezzati et al., 2021). Until recently, standard unsupervised clustering techniques have most commonly been used to identify subtypes using regional neuroimaging data (Habes et al., 2020). A recent meta-analysis of reported brain atrophy subtypes related to AD found studies consistently identified up to four subtypes: typical, limbic-predominant, hippocampal-sparing, and minimal atrophy (Ferreira et al., 2020). Additionally, they identified several common differences between subtypes including age at onset and at assessment, sex, years of education, APOE4 status, cognitive status, disease duration, and cerebrospinal fluid biomarker levels (Ferreira et al., 2020).

In this study, we have implemented an unsupervised learning clustering technique based on artificial neural networks called Self-Organizing Maps (SOM) (Kohonen, 1982/01) to identify patient sub-groups using volumetric MRI data. To our knowledge, this approach clustering technique of regional neuroimaging data has not previously been applied to identifying atrophy subtypes of AD. We used volumetric measures of nine brain regions of interest (ROIs) in amyloid positive patients with mild dementia enrolled in the placebo arm of EXPEDITION 3, a Phase 3 clinical trial of Solanezumab, an anti-amyloid monoclonal antibody binding soluble amyloid-β peptide. Differences in baseline characteristics between

clusters were assessed as well as differences in longitudinal cognitive performance using the ADAS-14 score, the primary outcome of EXPEDITION 3.

2. METHODS

2.1. Participants

We used data from the placebo arm the of the double-blind, placebo-controlled phase 3 trial of Solanezumab for mild AD (EXPEDITION 3) trial. EXPEDITION 3 was a phase 3 interventional clinical trial conducted between 2013 and 2017. Participants of the placebo arm received a placebo every 4 weeks for 76 weeks with an additional 4 weeks of assessments. The criteria for inclusion for male and female patients were to be between the ages of 55 and 90, to meet the National Institute of Neurological and Communicative Disorders and the Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for probable AD, and to have biomarker evidence consistent with the presence of amyloid at screening based on florbe-tapir PET scan or CSF. Exclusion criteria can be found elsewhere (Doody et al., 2014; Siemers et al., 2016). The clinical trial was approved by the by the ethics and institutional review boards of all sites. Participants provided written consent before trial participation.

For inclusion in this study, participants had baseline measures of demographics, neuropsychological tests, and volumetric magnetic resonance imaging. A subset of participants with longitudinal MRI were used to assess cluster differences in volumetric MRI changes.

2.2. Neuropsychological tests

The primary outcome measure for the trial was change in baseline in the Alzheimer's Disease Assessment Scale-Cognitive 14 item Subscore (ADAS-14, score range: 0–90, higher scores imply more impairment) (Rosen et al., 1984; Mohs et al., 1997). Secondary outcome measures include change from baseline on the following tests: Mini-Mental State Examination (MMSE, score range: 0–30, lower scores imply more impairment) (Folstein et al., 1975), Clinical Dementia Rating-Sum of Boxes (CDR-SB, score range: 0–18, higher scores imply more impairment) (Berg et al., 1988), Alzheimer's Disease Cooperative Study-Instrumental Activities of Daily Living (ADCS-iADL, score range: 0–78, lower scores imply more impairment) (Galasko et al., 2006), Functional Activities Questionnaire (FAQ, score range: 0–100, higher scores imply more impairment) (Pfeffer et al., 1982), and Integrated Alzheimer's Disease Rating Scale (iADRS, score range: 0–146, lower scores imply more impairment) (Wessels et al., 2015). The Geriatric Depression Scale (GDS, score range: 0–15, higher scores imply more impairment) (Sheikh and Yesavage, 1986) was used to assess depressive symptoms.

2.3. Magnetic resonance imaging

Participants had Volumetric MRI at screening visit. ROI-based volumetric analysis was performed using the FreeSurfer software (https://surfer.nmr.mgh.harvard.edu/). A detailed explanation of the procedure has previously been described (Schwarz et al., 2019). Available ROIs included: entorhinal cortex, hippocampus, inferior parietal lobe,

cingulate isthmus, lateral parietal lobe, medial temporal lobe, precuneus, prefrontal cortex, and superior temporal lobe. To account for differences in intracranial volume (ICV), ROIs volumes were adjusted (ROIa) using a residual correction approach such that $ROIa = ROI - B^{\star}(ICV - mean(ICV))$ where B is obtained as a regression coefficient when ROI volume is regressed against ICV.

2.4. Unsupervised clustering vis self-organizing maps

A multistep procedure was used for clustering of MRI volumetric data using Self-Organizing Maps, an artificial neural network approach trained using competitive learning. First, MRI data was normalized to the interval [0,1] and highly correlated regions removed. Second, principal component analysis was performed such that only components that increased the cumulative explained variance (CEV) by at least 10% were utilized in the clustering resulting in four principal components being used. Third, the SOM Toolbox uses the heuristic $d = ceil((5^{\star}N^{0.54321})^{1/2}) = 15$ (where N = 1041 is the number of participants) to determine the grid size such that the original data is projected onto a 15x15 grid of neurons using the SOM algorithm (Vatanen et al., 2015). The U-matrix and component planes associated with the SOM can be found in Supplementary Figs. 1 and 2. Finally, the SOM neurons were grouped using repeated K-mean clustering algorithm examining between 2 and 10 possible clusters. The optimal number of clusters was assessed using four metrics that include the Gap Statistic, Silhouette, Calinski-Harabbasz, and Davies-Bouldin values.

2.5. Statistical analysis

Statistical analysis of characteristic data was performed using MATLAB (version 2021a) and the SOM Toolbox 2.0 was used for implementing the self-organizing maps algorithm (Vatanen et al., 2015). All tests were two-tailed and the level of significance was taken at $\alpha = 0.05$. Characteristics among clusters were assessed using analysis of variance (ANOVA) for continuous variables and χ^2 tests for categorical variables. Post hoc analysis was performed using the least significant difference test. A repeated measures model was used to assess least square mean change from baseline of the ADAS-14 stratified by cluster at weeks 12, 28, 40, 52, 64, and 80. Spider plots were created of normalized neuropsychological tests and MRI volumes for each cluster. Plots were produced at the beginning and end of the study to assess change in each measure. Normalized values of the MMSE, iADRS, and ADL scores have been reversed such that an increase in normalized values corresponds to worse performance for all neuropsychological measures. Decrease in normalized MRI values corresponds with neurodegeneration. To investigate the stability of clusters as well as implications regarding severity, the SOM mapping trained on baseline data was used to predict cluster membership of individuals with follow-up data. An alluvial plot was used to visualize cluster membership at baseline and the follow-up visit.

3. RESULTS

3.1. Sample characteristics

The 1041 participants included in this study had an average age of 73.03 (SD = 7.81), had on average 13.67 (3.75) years of education, 58.6% were female, and 66.0% were APOE4

positive. In the entire population at baseline, the mean ADAS-14 was 29.65 (8.50), mean MMSE was 22.66 (2.88), mean CDR-SB was 3.91 (1.92), ADCS-iADL was 40.80 (11.54), FAQ was 10.57 (7.08), GDS was 1.65 (1.46) and IADRS was 105.77 (13.94). Table 1 summarizes sample characteristics of demographics, neuropsychological tests, and MRI regional volumes for the entire population and the clusters. Characteristics for the subgroup with longitudinal MRI can be found in Supplementary Table 1.

3.2. Unsupervised clustering

Different metric identified different optimal values of clusters. The Gap Statistics, Silhouette, and Calinski-Harabasz values favored a smaller number of clusters whereas the Davies-Bouldin value favored five or more clusters. Balancing these and noting abrupt changes in the Gap Statistic and Silhouette value at three clusters we identified three clusters as optimal (Supplementary Fig. 3).

3.3. Neuropsychological tests

ANOVA followed by post hoc analysis comparing baseline measures of neuropsychological tests was done. ANOVA comparing baseline neuropsychological measures between clusters indicated significant differences among all measures except GDS ($p = 0.384$). From post hoc analysis we found, for all measures, there to be significant differences at baseline between clusters except for Clusters 2 and 3 on the MMSE and between Clusters 1 and 2 on FAQ and ADCS-iADL. A summary of post hoc analysis results can be found in Supplementary Table 2. Spider plots show normalized test performance at baseline and follow-up in the subgroup (Fig. 1) where larger values corresponding to worse test performance.

Results from repeated measure models for ADAS-14, found between-cluster differences in the rate of change in ADAS-14 score during 80 weeks of trial (Fig. 1). The mean change from baseline to week 80 in ADAS-14 score for Cluster 1 was 4.28 (95% confidence interval [CI] = 3.60–4.96), for Cluster 2 was 10.14 (CI = 9.15–11.14), and for Cluster 3 was 7.90 (CI = 7.06–8.74). Significant between-group mean differences were observed between Clusters 1 and 2 (difference, 2.38; 95% CI, 1.85 to 2.91; $P < 0.001$), Clusters 1 and 3 (difference, 1.93; 95% CI, 1.41 to 2.44; $P < 0.001$), but not between Clusters 2 and 3 (difference, 0.45; 95% CI, −0.11 to 1.02; $P = 0.146$) in ADAS-14.

3.4. Baseline differences in volumetric MRI

Table 1 and Fig. 1 summarize baseline volumetric MRI measures and show significant differences in the pattern of ROI volumes between clusters. Cluster 1, the largest group (38.4% of the cohort), had the least neurodegenerative changes indicated by the largest ROI volumes. Cluster 2, the smallest group (25.8% of the cohort), exhibited little atrophy in ROI volumes in the entorhinal cortex, hippocampus, and medial temporal lobe but increased atrophy in the inferior and lateral parietal lobes, precuneus, and isthmus port of cingulate. Cluster 3 (35.7% of the cohort) had the lowest volumes of the entorhinal cortex, hippocampus, and medial temporal lobe of the three clusters, but slightly larger volumes than Cluster 2 for the inferior and lateral parietal lobes, precuneus, and isthmus port of cingulate. Spider plots of regional MRI volumes show smaller volumes for Clusters 2 and 3 compared to Cluster 1 (Fig. 1). ANOVA and post hoc analysis of longitudinal MRI measures

can be found in Supplementary Tables 3 and 4 Notably, asymmetries were observed that reflect regional differences in atrophy.

Additionally, using the SOM mapping trained on baseline data, we predicted cluster membership according to follow-up visit data. We found the majority of individuals remain in their baseline clusters, that is, 84.3%, 88.2%, and 91.0% or individuals remained in Clusters 1, 2, and 3, respectively, between Visits 1 and 22. Looking at transitions between clusters, we found 12.6% of individuals in Cluster 1 at baseline transitioned to Cluster 3 and 8.1% of individuals in Cluster 2 at baseline transitioned to Cluster 3. An alluvial plot of cluster memberships between Visit 1 and Visit 22 can be found in Supplementary Fig. 4.

4. Discussion

In this study, we explored whether data-driven SOM clustering of individuals exclusively based on their regional brain volumes would lead to identifying clinically different subtypes. We found three distinct MRI-based clusters: 1) least atrophy across all regions, 2) largest atrophy of the frontal and parietal lobes, and 3) largest atrophy of the temporal lobe. Cluster 1 had the lowest rate of cognitive decline based on ADAS-14, while Cluster 2 exhibited the highest rate of decline. Clusters were determined to largely be stable over the course of the clinical trial with the majority of transitions between clusters resulting in individuals moving towards the more atrophied Cluster 3. These two characteristics align with the proposed core dimensions of heterogeneity: typicality and severity (Ferreira et al., 2020).

Additionally, we showed the feasibility of a straightforward application of using SOM clustering to identify atrophy-related AD subtypes. Importantly, SOM-based clustering has several advantages. First, it is ideal for studies involving truly large datasets as a dimension reduction technique that preserves data topology and other key features (Villmann et al., 1997). Second, SOMs can be extended to supervised and semi-supervised models that account for both labeled and unlabeled data that can improve clustering performance (Braga and Bassani, 2018; Riese et al., 2020).

Identifying heterogeneity in AD neuroimaging patterns are largely based on two methods of hypothesis-driven and data-driven approaches. In hypothesis-driven method, classification is based on “a prior” atrophy subtypes such as medial temporal versus neocortical patterns (Habes et al., 2020). While this method provides important information on differential clinical, biomarkers and prognostic characteristics of subgroups, they are limited to previously discovered neuropathological subtypes and cannot provide information on whether the classification is representative of the larger population of patients. Data-driven neuroimaging methods commonly confirm the predefined neuropathological subtypes, while potentially identifying novel neuropathological patterns. Unsupervised clustering, such as SOM, is the most common way of subtyping neuroimaging patterns (Hastie, 2009).

The clusters identified in our study share similarities with those found in previous research, but also exhibit differences that may be related to our specific methodology and study population. For instance, one study applied robust collaborative clustering to MRI scans and identified three atrophy clusters (Toledo et al., 2022), one cluster of individuals with

dementia exhibited hippocampal sparing followed by another cluster characterized with limbic predominant atrophy. Our Cluster 2, characterized by frontal and parietal atrophy, differs from their hippocampal sparing cluster. This discrepancy might be due to the inclusion/exclusion criteria of the EXPEDITION 3 trial, which focused on individuals with mild dementia and amyloid positivity, potentially excluding certain atrophy patterns seen in broader populations.

Our results also show parallels with other work which identified typical and atypical subtypes including minimal atrophy, limbic-predominant, and hippocampal sparing patterns (Poulakis et al., 2018/05). However, our clustering approach, based solely on MRI volumetric data, may have captured different aspects of brain atrophy compared to their random forest-based method, which incorporated additional clinical variables. Finally, a third study combined clustering of subcortical and cortical volumes, for individuals across the AD continuum, with event-based modeling (Young et al., 2018).

Importantly, recent work has applied clustering approaches to neuroimaging data such as tau PET. For example, a recent study identified four distinct spatiotemporal trajectories of tau pathology in Alzheimer's disease (Vogel et al., 2021). Our MRI-based clusters do not directly map onto these tau-based subtypes as differences are expected as atrophy patterns and tau accumulation represent different aspects of the disease process. However, there are potential parallels. Their findings of distinct demographic and cognitive profiles associated with tau subtypes align with our observations of differential cognitive outcomes across MRI-based clusters. Their findings suggest that distinct neural network patterns may be associated with different disease progression trajectories, which aligns with our observations of heterogeneous atrophy patterns across clusters. Future research combining longitudinal MRI and tau PET data could elucidate the relationships between tau accumulation and atrophy across different AD subtypes.

To address potential entangling of phenotype differences with disease severity, we examined the temporal consistency of cluster classifications across visits. Analysis comparing initial cluster assignments to those at follow-up revealed high stability 88.1% maintaining their original cluster classification. Specifically, retention rates for clusters were 84.3% Cluster 1, 88.2% for Cluster 2, and 91.0% for Cluster 3. This high degree of consistency suggests that our identified clusters likely represent stable phenotypic differences rather than transient states of disease severity. The observed stability is particularly noteworthy for Cluster 3, which showed the highest retention rate. While some individuals transitioned between clusters (11.9% overall), these changes were relatively balanced and could reflect a combination of factors including disease progression or measurement variability. The overall consistency supports the robustness of our subtype classifications and their potential utility as meaningful disease phenotypes rather than merely reflecting varying levels of disease severity.

A key implication of this work is its potential impact on clinical trial design. By identifying distinct subtypes within a population already fitting clinical trial criteria, we provide a method for further stratification of participants. This approach could be particularly valuable when designing clinical trial that aim to enroll participants likely to show disease

progression in the placebo arm and demonstrate treatment effects in the intervention arm. The ability to identify and potentially predict different progression rates among subtypes could inform sample size calculations, enhance statistical power, and allow for pre-specified subgroup analyses in future trials. Additionally, stability of these subtypes over the course of a trial could be used as a basis for long-term stratification in longitudinal studies.

Several limitations of this study should be noted. First, this study was limited to individuals enrolled in a randomized control trial (RCT) who met certain enrollment criteria. As such, this study and its implications are limited to particular stages of AD progression. Second, this study did not have healthy individuals available to for use as a control group could be used for differentiating healthy and diseased aging as well as accounting for potential confounding variables and selection bias in the sample. Third, limited by data accessibility, no harmonization of MRI data was performed to account for different scanner types. Finally, we did not have access to other similar studies for external validation of our results. One potential issue with unsupervised clustering approaches is that variance captured by the identified subtypes may be related to different disease stages rather than to different disease expression patterns. This limitation has led other investigators to move toward disentangling AD heterogeneity through newer semi-supervised learning methods (Dong et al., 2016; Yang et al., 2021/12). Considering that we used data from a more homogenous group of individuals from an RCT who were all amyloid positive and in the same stage of disease, the latter limitation should be less of an issue.

Future directions of this work include applying this approach to additional datasets that cover the larger continuum of AD progression. Second, this work could be extended through implementation of novel semi-supervised and supervised SOM clustering that would aim to cluster disease effects in addition to clustering individuals (Braga and Bassani, 2018; Riese et al., 2020). Additionally, integrating our MRI-based clustering approach with tau PET data could provide a more comprehensive understanding of AD subtypes and their progression.

This work demonstrates the potential for advanced unsupervised learning methods to identify disease subtypes that were demonstrated to largely be stable over the course of this clinical trial with the majority of cluster transitions occurring towards Cluster 3. The subtypes show different patterns of cognitive decline, which might have implications in patient selection and design of interventional studies. Future studies are required to externally validate our findings and evaluate utility of clustering based on longitudinal imaging data, potentially incorporating multiple imaging modalities to capture the full spectrum of AD pathology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data used in these analyses are from the following Eli-Lilly trial: Progress of Mild Alzheimer’s Disease in Participants on Solanezumab Versus Placebo (EXPEDITION 3), ClinicalTrials.gov Identifier: NCT01900665.

This work was supported by grants from the National Institute in Aging (NIA K23 AG063993, Ezzati; NIA P01 AG003949, Lipton), the Alzheimer's Association (Ezzati, 2019-AACSF-641329), the Cure Alzheimer's Fund (Ezzati, Lipton), and the Leonard and Sylvia Marx Foundation (Lipton).

Disclosures

K.K.P has no disclosures.

B.T.N. has no disclosures.

R.B.L. receives research support from the following sources unrelated to this manuscript: 5U10 NS077308 (PI), R21 AG056920 (Investigator), 1RF1 AG057531 (Site PI), RF1 AG054548 (Investigator), 1RO1 AG048642 (Investigator), R56 AG057548 (Investigator), U01062370 (Investigator), RO1 AG060933 (Investigator), RO1 AG062622 (Investigator), 1UG3FD006795 (mPI), 1U24NS113847 (Investigator). He holds stock options in eNeura Therapeutics and Biohaven Holdings; serves as consultant, advisory board member, or has received honoraria from: Abbvie (Allergan), American Academy of Neurology, American Headache Society, Amgen, Avanir, Biohaven, Biovision, Boston Scientific, Dr. Reddy's (Promius), Electrocore, Eli Lilly, eNeura Therapeutics, Equinox, GlaxoSmithKline, Grifols, Lundbeck (Alder), Merck, Pernix, Pfizer, Supernus, Teva, Trigemina, Vector, Vedanta. He receives royalties from Wolff's Headache 7th and 8th Edition, Oxford Press University, 2009, Wiley and Informa.

E.G. receives a small royalty for commercial use of the Free and Cued Selective Reminding Test with Immediate Recall (FCSRT + IR). The test is available at no cost to researchers and clinicians. The Albert Einstein College of Medicines holds the copyright for the test.

A.E. serves as consultant, advisory board member, or has received honoraria from: PCORI Health Care Horizon Scanning System, GlaxoSmithKline, Mist Research, and Corium.

Data availability

The authors do not have permission to share data.

References

Berg L, Miller JP, Storandt M, et al. , 1988. Mild senile dementia of the alzheimer type: 2. Longitudinal assessment. Ann. Neurol 23 (5), 477–484. 10.1002/ana.410230509. [PubMed: 3389756]

Braga PHM, Bassani HF, 2018. A Semi-supervised Self-Organizing Map for Clustering and Classification, pp. 1–8.

Dong A, Honnorat N, Gaonkar B, Davatzikos C, 2016. CHIMERA: clustering of heterogeneous disease effects via distribution matching of imaging patterns. IEEE Trans. Med. Imag 35 (2), 612–621. 10.1109/tmi.2015.2487423.

Doody RS, Thomas RG, Farlow M, et al. , 2014. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. N. Engl. J. Med 370 (4), 311–321. 10.1056/NEJMoa1312889. [PubMed: 24450890]

Ezzati A, Abdulkadir A, Jack CR Jr., et al. , 2021. Predictive value of ATN biomarker profiles in estimating disease progression in Alzheimer's disease dementia. Alzheimer's Dementia 17 (11), 1855–1867. 10.1002/alz.12491.

Ferreira D, Nordberg A, Westman E, 2020. Biological subtypes of Alzheimer disease: a systematic review and meta-analysis. Neurology 94 (10), 436–448. 10.1212/wnl.0000000000009058. [PubMed: 32047067]

Folstein MF, Folstein SE, McHugh PR, 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res 12 (3), 189–198. 10.1016/0022-3956(75)90026-6. [PubMed: 1202204]

Galasko D, Bennett DA, Sano M, Marson D, Kaye J, Edland SD, 2006. ADCS Prevention Instrument Project: assessment of instrumental activities of daily living for community-dwelling elderly individuals in dementia prevention clinical trials. Alzheimer Dis. Assoc. Disord 20 (4 Suppl. 3), S152–S169. 10.1097/01.wad.0000213873.25053.2b. [PubMed: 17135809]

Gauthier S, Zhang H, Ng KP, Pascoal TA, Rosa-Neto P, 2018. Impact of the biological definition of Alzheimer's disease using amyloid, tau and neurodegeneration (ATN): what about the role of vascular changes, inflammation, Lewy body pathology? Transl. Neurodegener 7, 12. 10.1186/S40035-018-0117-9. [PubMed: 29876101]

Habes M, Grothe MJ, Tunc B, McMillan C, Wolk DA, Davatzikos C, 2020. Disentangling heterogeneity in Alzheimer's disease and related dementias using data-driven methods. Biol. Psychiatr 88 (1), 70–82. 10.1016/j.biopsych.2020.01.016.

Hastie TTRFJ, 2009. The elements of statistical learning. Data Mining, Inference, and Prediction, second ed. Springer.

Jack CR Jr., Bennett DA, Blennow K, et al. , 2018. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement 14 (4), 535–562. 10.1016/j.jalz.2018.02.018. [PubMed: 29653606]

Kohonen T, 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern 43 (1), 59–69. 10.1007/BF00337288.

Mohs RC, Knopman D, Petersen RC, et al. , 1997. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis. Assoc. Disord 11 (Suppl. 2), S13–S21. [PubMed: 9236948]

Pfeffer RI, Kurosaki TT, Harrah CH Jr., Chance JM, Filos S, 1982. Measurement of functional activities in older adults in the community. J. Gerontol 37 (3), 323–329. 10.1093/geronj/37.3.323. [PubMed: 7069156]

Poulakis K, Pereira JB, Mecocci P, et al. , 2018. Heterogeneous patterns of brain atrophy in Alzheimer's disease. Neurobiol. Aging 65, 98–108. 10.1016/j.neurobiolaging.2018.01.009. [PubMed: 29455029]

Riese FM, Keller S, Hinz S, 2020. Supervised and semi-supervised self-organizing maps for regression and classification focusing on hyperspectral data. Rem. Sens 12 (1), 7.

Rosen WG, Mohs RC, Davis KL, 1984. A new rating scale for Alzheimer's disease. Am. J. Psychiatr 141 (11), 1356–1364. 10.1176/ajp.141.11.1356. [PubMed: 6496779]

Schwarz AJ, Sundell KL, Charil A, et al. , 2019. Magnetic resonance imaging measures of brain atrophy from the EXPEDITION3 trial in mild Alzheimer's disease. Alzheimer's Dementia 5, 328–337. 10.1016/j.trci.2019.05.007.

Sheikh JI, Yesavage JA, 1986. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. Clin. Gerontol.: The Journal of Aging and Mental Health 5 (1–2), 165–173. 10.1300/J018v05n01_09.

Siemers ER, Sundell KL, Carlson C, et al. , 2016. Phase 3 solanezumab trials: secondary outcomes in mild Alzheimer's disease patients. Alzheimers Dement 12 (2), 110–120. 10.1016/j.jalz.2015.06.1893. [PubMed: 26238576]

Toledo JB, Liu H, Grothe MJ, et al. , 2022. Disentangling tau and brain atrophy cluster heterogeneity across the Alzheimer's disease continuum. Alzheimer's Dementia: Translational Research & Clinical Interventions 8 (1), e12305. 10.1002/trc2.12305. [PubMed: 35619830]

Vatanen T, Osmala M, Raiko T, et al. , 2015. Self-organization and missing values in SOM and GTM. Neurocomputing 147, 60–70. 10.1016/j.neucom.2014.02.061.

Villmann T, Der R, Herrmann M, Martinetz TM, 1997. Topology preservation in self-organizing feature maps: exact definition and measurement. IEEE Trans. Neural Network 8 (2), 256–266. 10.1109/72.557663.

Vogel JW, Young AL, Oxtoby NP, et al. , 2021. Four distinct trajectories of tau deposition identified in Alzheimer's disease. Nat. Med 27 (5), 871–881. 10.1038/S41591-021-01309-6. [PubMed: 33927414]

Wessels AM, Siemers ER, Yu P, et al. , 2015. A combined measure of cognition and function for clinical trials: the integrated Alzheimer's disease rating scale (iADRS). J Prev Alzheimers Dis 2 (4), 227–241. 10.14283/jpad.2015.82. [PubMed: 27019841]

Yang Z, Nasrallah IM, Shou H, et al. , 2021. A deep learning framework identifies dimensional representations of Alzheimer's Disease from brain structure. Nat. Commun 12 (1), 7065. 10.1038/s41467-021-26703-z. [PubMed: 34862382]

Young AL, Marinescu RV, Oxtoby NP, et al. , 2018. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. Nat. Commun 9 (1), 4273. 10.1038/s41467-018-05892-0. [PubMed: 30323170]

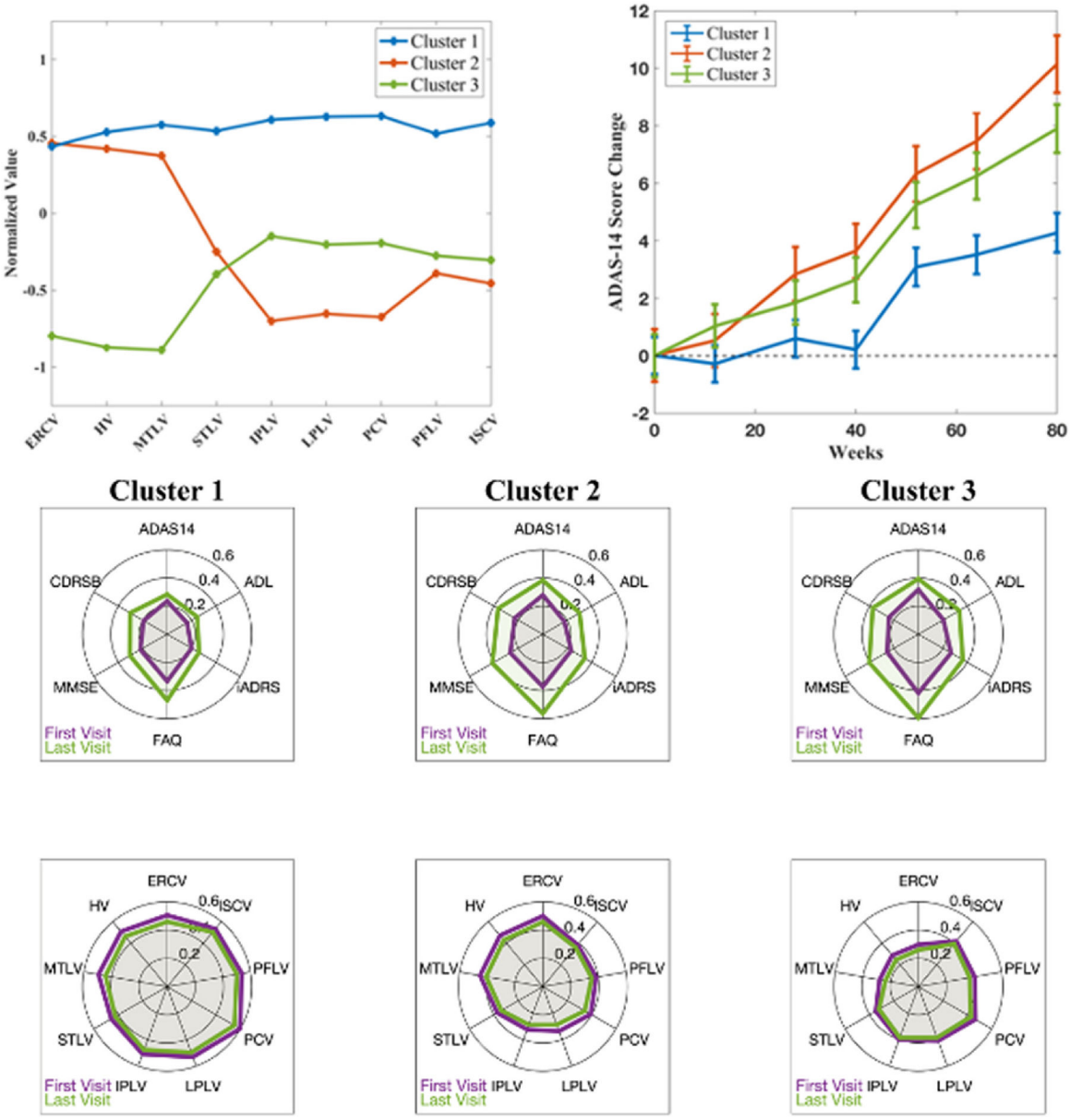

Fig. 1.

Regional MRI and cognitive profiles across three clusters. Top Left: Comparison of normalized regional volumetric MRI measures at baseline across clusters, revealing a gradient of neurodegeneration. Cluster 3 shows the most severe atrophy, particularly in the entorhinal cortex and hippocampus, while Cluster 1 exhibits the least overall atrophy. Top Right: Longitudinal changes in ADAS-cog14 scores, demonstrating differential rates of cognitive decline. Cluster 2 shows the steepest decline (increased score), Cluster 3 an intermediate rate, and Cluster 1 the slowest, delayed progression, suggesting varying disease trajectories. Bottom: Spider plots illustrating the multi-dimensional nature of cognitive decline and brain atrophy. (Top) Neuropsychological test scores at baseline (purple) and follow-up (green), normalized to a 0–1 scale. Cluster 1 consistently shows the best performance, with the least deterioration over time across all measures. (Bottom) Volumetric

MRI measures at baseline (purple) and follow-up (green), normalized to a 0–1 scale. Cluster 3 exhibits the most pronounced brain volume reductions, particularly in regions associated with early Alzheimer's disease. Note: For neuropsychological measures, higher normalized values indicate worse performance. For MRI volumes, decreases correspond to greater neurodegeneration. This visualization underscores the heterogeneity in disease presentation and progression across patient subgroups.

Abbreviations ADAS14, ADAS-cog14; ADL, instrumental subscale of the Alzheimer's Disease Cooperative Study Activities of Daily Living Inventory; iADRS, Integrated Alzheimer's Disease Rating; FAQ, Functional Activities Questionnaire; MMSE, Mini-Mental State Examination; CDRSB, Clinical Dementia Rating Sum of Boxes; ERCV, Entorhinal Cortex Volume; HV, Hippocampus Volume; MTLV, Medial Temporal Lobe Volume; STLV, Superior Temporal Lobe Volume; IPLV, Inferior Parietal Lobe Volume; LPLV, Lateral Parietal Lobe Volume; PCV, Precuneus Volume; PFLV, Prefrontal Lobe Volume; ISCV, Isthmus Port of Cingulate Volume. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1

Baseline characteristics of the entire sample and three distinct clusters, highlighting significant differences in demographic, genetic, cognitive, and functional measures. Cluster 1 showed the youngest mean age, highest cognitive function, and lowest impairment. Cluster 2 exhibited intermediate scores across most measures and had the highest proportion of APOE4 carriers. Cluster 3 demonstrated the oldest mean age and most severe cognitive and functional deficits. Significant differences were observed in age, APOE4 status, MRI volumetric measures, ADAS-14, MMSE, CDR-SB, ADCS-iADL, FAQ, and iADRS scores across clusters ($p < 0.05$). Depression scores and gender distribution remained comparable between clusters. Continuous variables were analyzed using ANOVA, and categorical variables with $\chi 2$ tests. Detailed post-hoc analyses are available in Supplementary Fig. 2.

N		**Entire Sample**	**Cluster 1**	**Cluster 2**	**Cluster 3**	**p**
		1041	**400**	**269**	**372**	
Demographics	**Age**	73.03 (7.81)	72.47 (7.29)	69.40 (8.62)	76.32 (6.25)	<0.001
	Sex, Female (%)	610 (58.6%)	222 (55.5%)	162 (60.2%)	226 (60.8%)	0.274
	Education, years	13.67 (3.75)	13.47 (3.65)	13.91 (3.90)	13.69 (3.76)	0.329
	APOE4 carrier (%)	665 (66.0%)	125 (32.6%)	116 (44.1%)	102 (28.3%)	<0.001
	ACHEI/Mem (%)	837 (80.4%)	301 (75.3%)	102 (37.9%)	317 (85.2%)	<0.001
Neuropsychological Tests	**ADAS-14, mean (SD)**	29.65 (8.50)	26.53 (7.52)	29.76 (8.73)	32.90 (8.11)	<0.001
	MMSE, mean (SD)	22.66 (2.88)	23.53 (2.79)	22.24 (2.88)	22.04 (2.76)	<0.001
	CDR-SB, mean (SD)	3.91 (1.92)	3.36 (1.67)	3.89 (1.75)	4.52 (2.09)	<0.001
	ADCS-iADL, mean (SD)	40.80 (11.54)	46.99 (7.54)	46.55 (7.00)	42.86 (8.75)	<0.001
	FAQ, mean (SD)	10.57 (7.08)	8.96 (7.02)	9.98 (6.63)	12.70 (6.93)	<0.001
	GDS, mean (SD)	1.65 (1.46)	1.68 (1.53)	1.72 (1.39)	1.57 (1.44)	0.384
	IADRS, mean (SD)	105.77 (13.94)	110.45 (12.40)	106.84 (13.16)	99.95 (13.99)	<0.001
MRI Volumes	**Entorhinal Cortex, cm^3**	2.66 (0.69)	2.96 (0.66)	2.97 (0.52)	2.11 (0.45)	<0.001
	Hippocampus, cm^3	5.96 (1.07)	6.53 (0.91)	6.41 (0.79)	5.03 (0.70)	<0.001
	Inferior Parietal Lobe, cm^3	21.00 (3.34)	23.04 (3.06)	18.67 (2.53)	20.51 (2.82)	<0.001
	Isthmus Cingulate, cm^3	3.93 (0.64)	4.31 (0.59)	3.65 (0.54)	3.74 (0.55)	<0.001
	Lateral Parietal Lobe, cm^3	59.86 (8.02)	6.49 (7.10)	54.63 (6.34)	58.23 (6.88)	<0.001
	Medial Temporal Lobe, cm^3	14.01 (2.37)	15.37 (2.08)	14.90 (1.63)	11.91 (1.45)	<0.001
	Precuneus, cm^3	15.46 (2.20)	16.85 (1.88)	13.97 (1.77)	15.03 (1.92)	<0.001
	Prefrontal Cortex, cm^3	49.97 (6.19)	53.18 (5.83)	47.56 (5.43)	48.27 (5.59)	<0.001
	Superior Temporal Lobe, cm^3	19.29 (2.67)	20.72 (2.58)	18.62 (2.31)	18.23 (2.32)	<0.001
	Intracranial Volume, cm^3	1496.23 (172.20)	1496.78 (177.98)	1495.26 (165.58)	1496.35 (171.02)	0.994

Note. Reported p-values are from individual ANOVAs. Abbreviations: ACHEI/Mem, use of acetylcholinesterase inhibitor and/or memantine; ADAS-14, Alzheimer's Disease Assessment Scale, 14-item Cognitive Subscale; MMSE, Mini-Mental State Examination; CDR-SB, Clinical Dementia Rating Sum of Boxes; ADCS-iADL, instrumental subscale of the Alzheimer's Disease Cooperative Study Activities of Daily Living Inventory; FAQ, Functional Activities Questionnaire; GDS, Geriatric Depression Scale; iADRS, Integrated Alzheimer's Disease Rating.