Nature Communications

https://doi.org/10.1038/s41467-026-72091-7

Article in Press

Coupled cross-sectional and longitudinal non-negative matrix factorization reveals dominant brain aging trajectories in 48,949 individuals

Received: 21 April 2025

Accepted: 6 April 2026

Cite this article as: Skampardoni, I., Erus, G., Nasrallah, I.M. *et al.* Coupled cross-sectional and longitudinal non-negative matrix factorization reveals dominant brain aging trajectories in 48,949 individuals. *Nat Commun* (2026). https://doi.org/10.1038/s41467-026-72091-7

Ioanna Skampardoni, Guray Erus, Ilya M. Nasrallah, Zhijian Yang, Michael R. Duggan, Keenan A. Walker, Alexander Getka, Kyunglok Baik, Randa Melhem, Sindhuja T. Govindarajan, Susan M. Resnick, Haochang Shou, Konstantina Nikita & Christos Davatzikos

Title

Coupled Cross-sectional and Longitudinal Non-Negative Matrix Factorization Reveals Dominant Brain Aging Trajectories in 48,949 Individuals

Author list

Ioanna Skampardoni[1,2], Guray Erus[1], Ilya M. Nasrallah[1,3], Zhijian Yang[1], Michael R. Duggan[4], Keenan A. Walker[4], Alexander Getka[1], Kyunglok Baik[1], Randa Melhem[1], Sindhuja T. Govindarajan[1], Susan M. Resnick[4], Haochang Shou[1,5], Konstantina Nikita[2], Christos Davatzikos[1*]

Affiliations

(1) AI2D Center for AI and Data Science for Integrated Diagnostics, University of Pennsylvania, Philadelphia, PA, USA, (2) School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece, (3) Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA, (4) Laboratory of Behavioral Neuroscience, National Institute on Aging, Baltimore, MD, USA, (5) Department of Biostatistics, Epidemiology, & Informatics, University of Pennsylvania, Philadelphia, PA, USA.

*Corresponding author: christos.davatzikos@pennmedicine.upenn.edu

Abstract

Machine learning can unravel heterogeneous patterns of brain aging and neurodegeneration, but existing methods offer limited insights into disease progression due to reliance on cross-sectional data. We introduce Coupled Cross-sectional and Longitudinal Non-negative Matrix Factorization (CCL-NMF) to capture dominant brain aging patterns by simultaneously leveraging cross-sectional and longitudinal neuroimaging data. CCL-NMF allows individuals to co-express multiple patterns, capturing mixed neuropathologic processes. Applied to neuroimaging data from 48,949 individuals from the harmonized iSTAGING study, CCL-NMF identifies seven distinct, reproducible, and biologically relevant neuroanatomical patterns. Subject-specific loading coefficients quantifying the individual expression of these patterns show distinct associations with cognition, genetic, and lifestyle factors. To support broader application, a regression-based tool was developed to estimate loadings in external cohorts without rerunning the full framework. By enabling individualized estimation of distinct brain aging patterns, these findings may improve risk assessment and therapeutic evaluation in neurodegenerative diseases. Although demonstrated using structural MRI, this framework is generalizable to other imaging modalities and biomarker types.

Introduction

Advances in machine learning (ML) and large-scale neuroimaging have enabled the extraction of informative biomarkers from cross-sectional (magnetic resonance imaging) MRI data, offering

valuable insights into brain aging and disease[1,2]. However, the heterogeneous and dynamic nature of brain changes in aging and neurodegeneration necessitates models that move beyond static, case-control frameworks. Integrating longitudinal data is essential to capture temporal progression and individual-specific trajectories. Previous approaches, such as Bayesian Latent Dirichlet Allocation[3] and SuStaIn (Subtype and Stage Inference)[4] have modeled latent atrophy patterns and disease stages from cross-sectional data but rely on discretization of continuous voxel data, predefined subtypes, and limited spatial resolution, potentially reducing sensitivity and oversimplifying disease complexity. Furthermore, both methods rely on cross-sectional data that limit their ability to capture individual-specific atrophy progression over time, and they parse heterogeneity directly in the patient domain, where disease-irrelevant confounding variability—such as neurodevelopmental differences—may obscure disease-relevant patterns. Weakly-supervised clustering techniques[5,6,7] address confounding by contrasting reference and target groups but often extract categorical subtype memberships, limiting their ability to capture the continuous and overlapping nature of brain aging. Recently, Surreal-GAN[8,9], a weakly-supervised, deep-representation learning method based on generative adversarial networks (GAN), has improved upon this by learning continuous low-dimensional representations (R-indices) of neuroimaging patterns; however, like prior models, it does not incorporate temporal information and therefore cannot capture dynamic progression.

In recent years, non-negative matrix factorization (NMF[10,11]), a data-driven approach in which a non-negative matrix is decomposed into two lower-rank non-negative matrices, has gained prominence as a robust technique for analyzing high-dimensional data across several fields, including text mining[12,13], speech processing[14,15], and neuroimaging[16,17,18]. NMF breaks complex data into additive building blocks, so that each observation can be expressed as a weighted combination of these parts. When applied to neuroimaging, these parts correspond to sparse, part-based representations that capture regions co-varying across individuals, aligning with established structural and functional networks. The decomposition simultaneously performs soft clustering and dimensionality reduction, providing an interpretable framework for uncovering hidden structure in large datasets. By capturing systematic covariation across individuals using cross-sectional data, NMF also enables the modeling of complex brain changes associated with development, aging, and disease. Despite these advances, the potential of NMF to capture patterns of pathological brain changes through the integration of longitudinal data remains largely untapped.

In this study, we introduce a methodology called Coupled Cross-sectional and Longitudinal NMF (CCL-NMF), which aims to identify distinct components that capture heterogeneous patterns of brain changes simultaneously from cross-sectional and longitudinal data through a joint optimization formulation. Building upon the methodological foundation of our preliminary model[19], we applied this approach to a large, diverse, harmonized cohort of 48,949 individuals from 12 neuroimaging studies across three continents.

Large cross-sectional datasets are widely available and enable the characterization of cumulative brain changes associated with aging and disease. However, they are limited by idiosyncratic effects and the absence of individualized baselines, necessitating population-level inferences. In contrast, longitudinal data offer a subject-specific view of neurobiological trajectories but remain comparatively scarce. To leverage the complementary strengths of both data types, CCL-NMF

employs a mutually constrained NMF framework that jointly factorizes cross-sectional and longitudinal inputs, each with different sample sizes, to extract shared components reflecting distinct patterns of brain alteration. Unlike models that impose rigid subtype membership, CCL-NMF allows individuals to co-express multiple components to varying degrees, a feature critical for capturing co-existing pathologies.

The proposed methodology is formulated in a general framework, enabling its application to heterogeneity analyses of any disease characterized by monotonic brain alterations (e.g., gradual brain atrophy or white matter hyperintensity (WMH) accumulation as measured by MRI or increasing deposition of neuropathologies such as amyloid and tau as measured by positron emission tomography). Here, we focus on applying this technique to study the heterogeneity of aging-related gray matter (GM) atrophy using regional volumes derived from T1-weighted (T1-w) MRI. First, we validate CCL-NMF on semi-synthetic data with predefined (ground truth) atrophy patterns and severity levels. Then, we apply it to a diverse, multi-cohort aging population, where it identifies seven dominant components of brain atrophy that are consistent cross-sectionally and longitudinally. The identified components correlate with Alzheimer's disease (AD) biomarkers, cognitive function, cardiovascular disease (CVD)-related factors, and cognitive decline, demonstrating the method's ability to disentangle aging- and disease-related neurodegenerative processes.

The present study improves our understanding of brain aging heterogeneity in several respects. First, it introduces an ML methodology for integrating cross-sectional and longitudinal neuroimaging measurements within a unified framework, capitalizing on the advantages of each data type. Second, when applied to a large multi-cohort dataset, it reveals key aging- and disease-related trajectories by capturing seven reproducible and clinically relevant brain atrophy patterns. Comparisons with another state-of-the-art deep learning (DL) model demonstrate that the CCL-NMF components provide improved predictive performance for biomarkers and clinical variables, including amyloid, tau, cognitive decline, hypertension, obesity, and Apolipoprotein E (*APOE*) status, as well as for disease progression, thereby refining our grasp of brain aging pathways. Finally, the study offers a practical approach for researchers to quantify the expression of these seven brain atrophy patterns in their datasets through simplified, readily applicable models.

Results

An overview of the CCL-NMF framework is shown in Fig. 1, illustrating its application to structural MRI data to identify components capturing patterns of GM atrophy. The model jointly factorizes two complementary sources of brain imaging data: the C-map, which captures cross-sectional deviations of the aging population from a middle-aged reference group, and the L-map, which encodes individualized atrophy rates estimated from longitudinal measurements. This coupled decomposition yields a shared set of latent components (dictionary) representing distinct brain atrophy patterns and assigns separate subject-specific loading coefficients for each data type. Full methodological details, including map estimation and the joint NMF formulation, are provided in the Methods. Although generalizable to any condition involving monotonic brain alterations, the framework is here demonstrated in the context of aging-related GM atrophy. The biological and predictive utility of the CCL-NMF representation is also summarized in Fig. 1. Note that the figure does not depict the semi-synthetic validation experiments, which are described below.

Throughout this study, components denote the latent spatial factors identified by CCL-NMF, patterns refer to the neuroanatomical brain-change configurations captured by these components, and trajectories denote subject-specific longitudinal atrophy patterns.

Model validation using semi-synthetic dataset

Volumes of 119 predefined GM regions of interest (ROIs) (Supplementary Table 1) were used as input features in all analyses. Method validation was first performed using semi-synthetic data with known ground truth. This dataset was generated by simulating five distinct atrophy patterns on ROI volumes (Supplementary Table 2) derived from a cohort of cognitively normal (CN) individuals. A normative model was trained on the original ROI values of 903 CN individuals ($S1_{syn}$, the reference population) and applied to 3,614 individuals with simulated atrophy ($S2_{syn}$, the target population) to estimate cross-sectional deviations. From this group, the 1,365 individuals showing the largest deviations ($S3_{syn}$) were selected to construct the C-map. A subset of 300 individuals from $S3_{syn}$ with simulated longitudinal measurements was used to construct the L-map, simulating the common scenario in which only a subset of participants has longitudinal data. Details of the semi-synthetic data generation and experimental setup are provided in the Methods. The alignment between the estimated patterns and the ground truth was evaluated across three NMF models (Fig. 2):

1) C-NMF: NMF model using solely the C-map (number of subjects in C-map N_C=1,365; synthetic mean age 77.11±11.31 years, 38.24% females, patterns: frontal: 319 (23%), occipital: 228 (17%), parietal: 239 (18%), subcortical: 307 (22%), temporal: 272 (20%)).

2) L-NMF: NMF model using solely the L-map (number of subjects in L-map N_L=300; synthetic mean baseline age 76.82±10.92 years, 40% females, patterns: frontal: 83 (28%), occipital: 49 (16%), parietal: 50 (17%), subcortical: 54 (18%), temporal: 64 (21%)).

3) CCL-NMF: NMF model using both C-map (N_C=1,365) and L-map (N_L=300).

The joint NMF model using C-map and L-map together can identify the simulated patterns with higher accuracy (Fig. 2). The norm of the divergence matrix, which quantifies overall disagreement between predicted and simulated ground truth patterns, was smaller for the joint model (n=1.39) compared to the model using the C-map (n=1.67) or the L-map (n=1.43) individually. While longitudinal data provide a more direct measure of brain change, their limited availability necessitates balancing their influence relative to cross-sectional data. Weighting the contributions of the two data types according to their sample size ratio yielded optimal performance (Supplementary Fig. 1). This weighting is controlled by a coefficient, α, which accounts for the unequal sample sizes between cross-sectional and longitudinal data in the joint decomposition (see Methods, Eq. (5)).

Application of CCL-NMF to the iSTAGING dataset

Having validated the model on semi-synthetic data, we next applied CCL-NMF to real-world neuroimaging data from the iSTAGING (imaging-based coordinate SysTem for AGIng and NeurodeGenerative diseases)[8,20,21] consortium to evaluate its biological relevance and practical utility. To mirror the structure used in the semi-synthetic experiments, analogous notation was adopted in the real-world setting: $S1_{real}$ denotes the reference population, $S2_{real}$ the full target aging population, and $S3_{real}$ the high-deviation subset used to construct the C-map. A normative model was trained on 119 GM ROI volumes from 977 CN, middle-aged individuals ($S1_{real}$) and applied to a target population of 48,949 individuals aged 50 years and older ($S2_{real}$). From this cohort, 13,950 individuals exhibiting the most pronounced deviations from the normative model ($S3_{real}$) were selected to construct the C-map. A subset of 1,063 individuals within $S3_{real}$ with longitudinal MRI data was used to estimate regional atrophy rates and generate the L-map. Demographic and clinical characteristics of the samples used to construct the C- and L-maps are detailed in the Methods and in Supplementary Tables 3 and 4.

Validation of the C-map

Before applying the joint NMF, we first evaluated the stability and subject specificity of the derived C-map. Because this step serves to ensure that the cross-sectional deviations provide a reliable representation of individual anatomical variation, validation emphasized robustness and consistency rather than conventional goodness-of-fit metrics. Specifically, the cross-sectional deviations exhibited high within-subject consistency across longitudinal follow-up scans, preserved subject-specific multivariate patterns of regional change over time, and remained stable across independent normative model initializations. Permutation-based null experiments further confirmed that this stability reflects genuine subject-level anatomical signal rather than modeling artifacts. These results indicate that the C-map captures stable individual deviation patterns suitable for downstream joint factorization. Full quantitative results, together with complementary analyses evaluating diagnostic-group sensitivity of the cross-sectional deviations, are provided in Supplementary Method 1, Supplementary Notes, and Supplementary Fig. 2.

CCL-NMF identifies seven distinct components of brain atrophy

The CCL-NMF was run with the number of components *K* ranging from 2 to 15. Split-half reproducibility index[17] and sparsity[22] were used to determine the optimal number of components (Fig. 3). Sparsity increased with higher *K*. The split-half reproducibility analysis revealed a declining trend in reproducibility as the number of components increased, with a peak occurring at *K*=3 and 7. *K*=7 was selected for subsequent analyses because of its higher sparsity than the *K*=3 solution. Supplementary Method 2 and Supplementary Fig. 3 provide more details about the metrics. Additional reproducibility analyses, including independent runs with different random initializations, bootstrap resampling of subjects, and reruns of the full pipeline with independently trained normative models, are presented in Supplementary Method 3 and Supplementary Figs. 4-6.

CCL-NMF identified seven distinct components of GM atrophy, described as:

- CCL-NMF1 captures atrophy primarily in portions of the basal ganglia (including the putamen and caudate), as well as in the orbital gyrus, gyrus rectus, and subcallosal area.
- CCL-NMF2 involves atrophy primarily in the medial temporal lobe, temporal pole, temporal gyrus, and fusiform gyrus.
- CCL-NMF3 represents atrophy in the inferior frontal gyrus, occipital gyrus, and part of the temporal gyrus.
- CCL-NMF4 exhibits medial frontoparietal atrophy, including the superior and middle frontal gyri, precuneus, middle and posterior cingulate gyri, supplementary motor cortex, and superior parietal lobule.
- CCL-NMF5 is characterized by perisylvian atrophy (insula, frontal and central opercula, and planum polare), and by atrophy in the anterior cingulate gyrus.
- CCL-NMF6 captures atrophy primarily in the basal ganglia, including the accumbens area, pallidum, and thalamus.
- CCL-NMF7 captures atrophy in the cerebellum and the medial occipital lobe, including the cuneus, calcarine cortex, and lingual gyrus.

Figure 4 shows brain maps of these seven components, which exhibit sparse, orthogonal, and bilaterally symmetrical atrophy patterns.

To assess the contribution of longitudinal data to the learned dictionary, while acknowledging its smaller sample size relative to the cross-sectional dataset, we compared the CCL-NMF components with those derived from separate NMF decompositions of the C-map and L-map matrices (Supplementary Method 4). Component dictionaries were aligned using the Hungarian[23] algorithm, and similarity between corresponding components was quantified using cosine similarity. The CCL-NMF dictionary showed stronger correspondence with the longitudinal-only dictionary (cosine similarity: 0.84 ± 0.14) than with the cross-sectional-only dictionary (cosine similarity: 0.66 ± 0.18), indicating that longitudinal information contributes substantially to the learned patterns despite the smaller number of subjects with longitudinal measurements.

Longitudinal loadings capture progressive brain atrophy patterns

A key strength of the CCL-NMF framework lies in its integration of longitudinal data, enabling the refinement of the derived components through the incorporation of subject-specific atrophy trajectories. To evaluate the relevance of longitudinal loadings, rates of regional volume change were compared between individuals with high and low loadings across components. Subjects were randomly sampled from the top and bottom 10th percentiles of each component's longitudinal loading distribution. Figure 5 shows representative ROIs for selected CCL-NMF components, with fitted regression lines illustrating atrophy trajectories in individuals with low (blue squares) and high (red circles) longitudinal loadings.

Individuals with higher loadings exhibited steeper volume decline in corresponding regions, supporting the ability of CCL-NMF to differentiate distinct atrophy trajectories. For instance, CCL-NMF2, which primarily captures medial temporal atrophy, corresponded to greater hippocampal volume decline in individuals with high longitudinal loadings. Similarly, CCL-NMF5, reflecting perisylvian atrophy, showed accelerated reduction in the insular cortex. CCL-NMF4, characterized by medial frontoparietal atrophy, was linked to a steeper decline in precuneus volume among individuals with higher loadings. These findings illustrate that longitudinal loadings reflect distinct regional atrophy trajectories over time. Unlike conventional cross-sectional approaches that infer brain aging patterns from static baseline observations, CCL-NMF integrates dynamic trajectories, providing a more comprehensive characterization of neurodegenerative progression.

Associations of cross-sectional loadings with cognitive, clinical, biomarker, and disease progression measures

To assess the biological relevance of the brain atrophy components, we examined associations between their subject-specific cross-sectional loading coefficients and a range of clinical features, cognitive performance metrics, AD biomarkers, *APOE* ε4 allele (*APOE4*) status, and measures of cognitive impairment progression (Fig. 6). The components exhibited distinct profiles across these domains. CCL-NMF2, characterized by medial temporal lobe atrophy, showed the strongest associations with AD pathology, cognitive decline, and progression from mild cognitive impairment (MCI) to AD. CCL-NMF5, defined by prominent perisylvian atrophy, showed strong links to CVD-related factors, particularly obesity, and hypertension, as well as WMH. Similarly, CCL-NMF6, defined by basal ganglia atrophy, was significantly associated with elevated WMH and obesity.

CCL-NMF3, defined by inferior frontal, occipital, and temporal atrophy, was closely related to advanced age and progression from CN to MCI. CCL-NMF3 also showed moderate associations with amyloid positivity and cognitive decline (although less pronounced than those for CCL-NMF2), and with WMH (albeit less strongly than for CCL-NMF5). This component exhibited higher and more broadly distributed loadings across the aging population compared to the loadings of other components (Supplementary Fig. 7), consistent with the hypothesis that it may reflect general aging effects rather than a specific pathological process. CCL-NMF4 displayed a unique association with tau pathology but showed no significant relationships with other AD biomarkers,

cognitive decline, or CVD-related factors. However, the small tau sample size limits definitive conclusions. Lastly, CCL-NMF7 was not significantly associated with AD biomarkers, CVD-related factors, or cognitive impairment progression. Among the seven components, CCL-NMF4 and CCL-NMF7 exhibited negative associations with age. These associations indicate that the expression of these patterns decreases with increasing age.

Comparison with Surreal-GAN

The CCL-NMF components align with, yet extend beyond, the five atrophy dimensions reported by Yang et al.[8], who applied a different methodology (Surreal-GAN, Supplementary Method 6) to the same dataset (Supplementary Fig. 8). Key differences between these two representations are evident in CCL-NMF6, characterized by atrophy in the accumbens area, and CCL-NMF7, defined by atrophy in the cerebellum and medial occipital lobe. These patterns were not captured as distinct dimensions in the Surreal-GAN representation.

To evaluate the additional information captured by the CCL-NMF representation compared to Yang et al.'s approach, predictive models were developed using CCL-NMF loading coefficients, Surreal-GAN R-indices, and a combination of both as predictors for various biomarkers and clinical variables, including amyloid, tau, hypertension, obesity, and *APOE4* status, and for disease progression. As shown in Fig. 7, both R-indices and CCL-NMF loadings improved predictive performance relative to demographic-only models, with CCL-NMF loadings consistently providing the greatest gains across all outcomes. This suggests that the longitudinal information embedded in CCL-NMF yields components that are more neuropathologically relevant than those derived from the cross-sectional Surreal-GAN approach. We further validated these improvements through formal statistical comparisons. Permutation tests across folds confirmed consistent gains in discrimination and explained variance, though these did not reach conventional significance thresholds. In contrast, likelihood ratio tests (LRTs) demonstrated significant improvements in model fit when adding CCL-NMF loadings compared to demographics-only and R-indices models across multiple outcomes (Supplementary Method 7 and Supplementary Tables 5–9). Adding R-indices to models already containing CCL-NMF loadings provided minimal further benefit.

Estimation of CCL-NMF loadings in external datasets

Although CCL-NMF is designed for discovery and establishes components using a large, harmonized dataset, re-deriving these components in new datasets can be computationally intensive. This is due to the need for a suitable reference population, harmonization procedures, and implementation of the full CCL-NMF pipeline described in the Methods. To facilitate broader use of these brain atrophy components, we trained regression models to estimate CCL-NMF loading coefficients directly from GM ROI volumes, age, sex, and intracranial volume (ICV^{24}). Notably, these ROI inputs were not harmonized, thereby eliminating the need for reference-based

harmonization when estimating loadings in external datasets, an essential step during discovery but often impractical in real-world applications, including single-subject scenarios.

Figure 8 illustrates the Spearman correlations between original (uppercase) and approximated (lowercase) CCL-NMF loadings. The correlations were high, ranging from 0.8 to 0.93 for cross-sectional and from 0.9 to 0.97 for longitudinal loadings. The cross-sectional loadings showed slightly lower correlations (diagonal elements in Fig. 8A) than longitudinal ones (diagonal elements in Fig. 8B). This difference likely reflects the L-map's subject-specific trajectories, which are less sensitive to harmonization variability. Additionally, the use of ROI volume deviations rather than the actual ROI volumes as inputs in the CCL-NMF framework may contribute to this discrepancy. Supplementary Fig. 9 presents the correlation matrices within original and approximated cross-sectional and longitudinal loadings. The preservation of cross-component correlation structure indicates that the approximated loadings not only recover individual component expression but also retain inter-component relationships.

In addition to correlation analyses, we validated the regression-based loadings in terms of predictive performance across multiple clinically relevant outcomes, including both binary (e.g., *APOE4* status, amyloid, tau, obesity, hypertension) and time-to-event (MCI-to-AD progression) measures. Predictive accuracy and model fit were highly comparable to those obtained with the original loadings, with no significant differences detected by permutation testing (Supplementary Method 8 and Supplementary Tables 10-11). Finally, Supplementary Fig. 10 shows that association tests based on approximated cross-sectional loadings replicated the results from Fig. 6, further supporting their reliability.

Together, these findings support the use of pre-trained regression models as a practical and reliable alternative to full CCL-NMF recomputation. Loading coefficient estimation is available via the NiChart (Neuro Imaging Chart of AI-based Imaging Biomarkers) platform (https://cloud.neuroimagingchart.com/).

Using this regression-based estimator, we further evaluated the generalizability of the learned CCL-NMF components in an independent AD-focused cohort (ADSP: https://adsp.niagads.org/) under a strict out-of-sample setting; these analyses are presented in Supplementary Fig. 11.

Discussion

Brain aging is highly heterogeneous and is shaped by genetic, lifestyle, and environmental factors that can contribute to neuropathological progression. MRI, a widely accessible imaging modality, enables detailed assessment of macroscopic neurodegeneration such as brain atrophy and small-vessel ischemic changes. While MRI does not directly capture specific neuropathologies, ML techniques can identify patterns of neurodegeneration that reflect these biological processes[25,26,27,28]. Here, such patterns are extracted using a coupled cross-sectional and longitudinal NMF framework (CCL-NMF) to define a succinct yet comprehensive neuroanatomical coordinate system that captures the multidimensional variability of brain aging and disease-related neurodegenerative processes. By integrating both cumulative cross-sectional deviations and subject-specific longitudinal trajectories, the model captures progressive structural alterations over time and delineates dominant atrophy patterns that underpin the neuroanatomical coordinate space. Importantly, the framework does not assign individuals to discrete subtypes but instead estimates continuous expression levels across multiple components, allowing for the simultaneous representation of overlapping atrophy patterns. Although tensor-based methods can capture more complex temporal dynamics, NMF offers a balance of interpretability, robustness to irregular follow-up, and scalability in large, heterogeneous cohorts. Building on prior frameworks such as SuStaIn, HYDRA, CHIMERA, and Surreal-GAN, which have substantially advanced the modeling of disease heterogeneity, CCL-NMF extends disease heterogeneity modeling by explicitly incorporating longitudinal information at scale.

Leveraging data from the large, diverse, multi-cohort iSTAGING consortium (predominantly cognitively normal participants) and using the CCL-NMF methodology, we mapped the neuroanatomical heterogeneity of brain aging. Seven distinct and reproducible brain atrophy components were identified. While alternative values of K can yield distinct but valid decompositions, the seven-component solution was selected based on complementary criteria of reproducibility and sparsity, and here we focus on these dominant, stable patterns. Among these, CCL-NMF2, which primarily captured medial temporal lobe atrophy, displayed characteristics indicative of Alzheimer's disease, such as a strong link with amyloid and tau deposition, *APOE4*, cognitive decline, and clinical progression from MCI to AD, aligning with previous findings[29,30,31]. CCL-NMF3, which was strongly associated with advanced age and broadly expressed across the aging population, exhibited moderate correlations with amyloid deposition, cognitive decline, and WMH. These findings suggest that CCL-NMF3 may predominantly reflect aging-related effects rather than a distinct pathological process. CCL-NMF5, marked by perisylvian atrophy, demonstrated strong associations with CVD pathology, consistent with prior literature[8]. The strength of associations varied across components and outcomes, with certain components (e.g., CCL-NMF2) showing robust disease-aligned relationships consistent with Alzheimer’s pathology and progression, whereas others exhibited more graded or weaker aging-related effects, reflecting the heterogeneous and multifactorial nature of brain aging and neurodegeneration.

CCL-NMF1, which primarily captured orbitofrontal atrophy, was associated with both obesity and hypertension, aligning with prior studies linking the orbitofrontal cortex to food-related behaviors[32,33]. Interestingly, while CCL-NMF5 and CCL-NMF6 were also associated with hypertension and obesity, they were strongly associated with WMH, whereas CCL-NMF1 was not. These observations point to two potentially distinct pathways linking obesity and hypertension

to neurodegeneration: one driven by orbitofrontal atrophy without WMH involvement, and another with broader atrophy patterns associated with high WMH load.

CCL-NMF4 and CCL-NMF7 exhibited negative associations with age. Importantly, this does not imply preservation or increases in regional brain volume; rather, these patterns are expressed less frequently among older individuals. Several factors may contribute. Selection bias is common in multi-cohort aging studies, as older participants who remain are typically healthier and may under-express certain atrophy patterns. Component overlap may also play a role, since CCL-NMF components are not orthogonal, and atrophy reflected in CCL-NMF4 or CCL-NMF7 may be captured by other, more dominant components that increase with age. The negative association of CCL-NMF4 loadings with WMH further suggests processes partly independent of vascular pathology, though this interpretation is tentative. Moreover, because most contributing cohorts are AD-focused, even cognitively normal participants may be selectively free of AD but not of other neurological or systemic conditions. Finally, with only AD- and vascular biomarkers available, we cannot assess additional biological processes that may underlie these patterns. Interpretations of these components should therefore be cautious, and future work incorporating complementary markers, including genetic and proteomic data, will be essential for refining their characterization.

To contextualize our findings, we tested the CCL-NMF representation of brain aging against Surreal-GAN. These methods are fundamentally distinct: Surreal-GAN uses a GAN-based weakly-supervised clustering approach, while CCL-NMF combines a DL-based normative model with longitudinal change maps to derive heterogeneity dimensions via NMF. A key distinction is that Surreal-GAN uses baseline cross-sectional measures, whereas CCL-NMF jointly models cross-sectional and longitudinal inputs, allowing each to inform the other. This joint modeling addresses a limitation of purely cross-sectional approaches by ensuring consistency between neurodegeneration patterns inferred from cross-sectional data and those observed through longitudinal progression. Both approaches were applied to the iSTAGING dataset to investigate aging-related brain atrophy.

The CCL-NMF components largely aligned with the five atrophy dimensions identified by Surreal-GAN, while providing an expanded representation. Predictive models incorporating CCL-NMF loadings consistently outperformed those using R-indices, underscoring the richer representation provided by CCL-NMF. Importantly, these improvements were not only qualitative but also validated by formal statistical comparisons: LRTs consistently showed that adding CCL-NMF loadings significantly improved model fit relative to demographics-only and R-indices models, even when permutation tests across folds did not always reach conventional significance. These findings emphasize that integrating longitudinal information yields measurable, statistically supported improvements in heterogeneity modeling. By capturing temporal dynamics, CCL-NMF extends beyond static cross-sectional models to enable a more nuanced understanding of brain aging and related pathologies.

To further contextualize our findings, we evaluated the generalizability of the learned CCL-NMF components in an independent Alzheimer's disease-focused cohort using out-of-sample loading estimation. The results showed overall consistency in key patterns of association and biological interpretation, supporting the robustness and external generalizability of the derived brain aging

axes while naturally reflecting cohort-specific differences in disease enrichment, age distribution, and variability of clinical and biomarker measures.

Our ML framework operates in two key phases: first, the C- and L-maps are estimated; second, a mutually constrained NMF is applied. This approach yields a shared dictionary across the two data types, along with corresponding loading coefficients for each data type. Although the framework incorporates well-established tools at various stages, implementing these procedures may present challenges for new users. To enhance accessibility and practical application, regression models were developed that accurately predict the loading coefficients derived from the CCL-NMF model. This simplification allows users to estimate these coefficients directly from their own datasets using a publicly available web-based platform without needing expertise in our two-step framework. Importantly, these regression models are built on raw data, so they do not require new users to harmonize their data before using them. By bridging the gap between model complexity and user-friendly application, CCL-NMF facilitates broader adoption of the model's predictive capabilities in diverse research settings.

Building on the widely used NMF framework, CCL-NMF provides a robust and flexible approach for disentangling heterogeneous brain changes by jointly modeling cross-sectional and longitudinal neuroimaging data. The integration of these complementary data types enables the identification of seven reproducible and clinically meaningful components of atrophy, each capturing a distinct aging-related trajectory, in a large, harmonized, multi-cohort dataset of 48,949 individuals from the iSTAGING consortium. In contrast to models that impose rigid subtype membership, CCL-NMF allows individuals to co-express multiple patterns, better reflecting the complexity of overlapping neuropathologies. By quantifying individualized expression levels across components, the method supports the development of personalized therapeutic strategies. Compared to a state-of-the-art DL model (Surreal-GAN), CCL-NMF demonstrated improved predictive performance for a range of biomarkers and clinical outcomes, thereby refining the characterization of aging-related neurodegeneration. While the current application focuses on aging-related atrophy, the framework is generalizable to other brain disorders characterized by monotonic progression. Finally, the inclusion of regression-based tools enables estimation of component loadings in new datasets without re-deriving the model or harmonizing the data, promoting broader adoption across diverse research settings.

While the model effectively captures individualized patterns of brain atrophy by integrating longitudinal data, its reliance on temporal information poses challenges in settings where such data are limited. In this study, the disproportionately smaller longitudinal sample may have constrained the influence of dynamic information, despite efforts to balance contributions through weighting. Another methodological consideration is the non-negativity constraint inherent to NMF. While this promotes part-based representations and interpretability, it requires single-signed inputs. In practice, this entails zeroing or sign-flipping mixed-signed values, which may reduce sensitivity to biological processes involving transient volumetric increases or compensatory effects. The present framework is therefore formulated to capture dominant progressive processes, such as regional atrophy, cerebrospinal fluid expansion, and WMH accumulation, that characterize brain aging and neurodegeneration. As such, non-negativity should be viewed as a modeling choice aligned with this objective rather than as a comprehensive representation of all possible brain changes.

A further limitation is the use of ROI-level features, which may obscure localized or subregional effects due to spatial averaging. Applying this framework at the voxel-level could increase anatomical specificity and enable the detection of more granular neuroanatomical changes.

In addition to these methodological aspects, sample size is another constraint. The relatively modest training set for the normative model is on the lower end for DL approaches and may restrict generalizability. We nevertheless selected the autoencoder for its ability to capture nonlinear inter-regional dependencies and higher-order structure in the data, features that univariate normative models cannot represent. Still, multivariate non-DL strategies may be more suitable in smaller samples and should be evaluated in future work.

The reference cohort used for normative modeling is restricted to cognitively normal midlife individuals without major cardiometabolic or vascular risk factors; therefore, the learned variability reflects a midlife reference manifold and may include undetected preclinical neurodegenerative variability rather than full lifespan normative variation. In this context, some biological interpretations remain tentative. Although CCL-NMF components show clear associations with Alzheimer's disease and vascular risk markers, the underpinnings of patterns such as CCL-NMF4 and CCL-NMF7, both negatively associated with age, are not fully understood. These findings highlight the need for complementary biological markers, including genetic and proteomic markers, to refine characterization. Moreover, while the regression-based estimator was trained on unharmonized ROI data to approximate site- and scanner-related variability, its stability under explicit scanner drift or in fully independent cohorts remains to be established. The contributing cohorts are predominantly of European ancestry, enriched for cognitively normal research participants, and not designed to represent clinical populations; accordingly, caution is warranted when extrapolating these population-level aging patterns to more diverse or clinically ascertained groups.

In conclusion, this study introduced an ML-based approach designed to disentangle the heterogeneity of brain aging by leveraging both cross-sectional and longitudinal data. Applied to structural MRI data from a large aging cohort, the method identified distinct, reproducible, and clinically relevant components associated with brain aging. This approach enhances traditional cross-sectional methods by integrating temporal dynamics, enabling more nuanced insights into the progressive nature of complex neurobiological processes underlying aging. It may also inform personalized interventions, refining therapeutic strategies for brain aging and neurodegenerative diseases.

Methods

CCL-NMF framework

CCL-NMF addresses brain aging heterogeneity by decomposing brain changes into distinct components using NMF. These components represent coordinated patterns of brain alteration that are likely linked to underlying neuropathologic processes. Specifically, a mutually constrained NMF framework is introduced that integrates two complementary sources of information: maps representing cross-sectional and longitudinal brain changes, referred to as the C-map and L-map, respectively. The C-map captures long-term aging-related deviations by comparing individuals to a population-based normative reference, while the L-map reflects dynamic brain changes over time at the individual level. Importantly, NMF is applied to these model-derived maps rather than directly to raw anatomical measurements. Using raw features would primarily cluster individuals according to sources of inter-individual variability unrelated to aging or neurodegeneration (e.g., sex, head size, or developmental differences), thereby reducing biological interpretability. In contrast, deriving deviation and change rate maps filters out such nuisance variation in a multivariate fashion, ensuring that the decomposition captures aging- and disease-related processes, rather than normative anatomical covariance. The joint NMF approach further identifies components shared by cross-sectional and longitudinal maps based on the assumption that aging- or disease-related effects inferred from cross-sectional deviations should align with the temporal dynamics observed in longitudinal data. The model also estimates corresponding loading coefficients representing the degree to which each subject expresses each component, by jointly optimizing the reconstruction of both maps. In this way, the model captures the interplay between static and dynamic aspects of brain alterations (Fig. 9).

The C-map is derived using a normative modeling approach trained on a middle-aged, cognitively normal reference cohort without major cardiometabolic or vascular risk factors, intended to approximate brain variability prior to substantial age-related neuropathology. This reference population is assumed to be relatively less affected by overt neuropathologic processes that typically emerge beyond midlife[33], although preclinical pathological changes may already be present. In this study, “normative modeling” is used in a reference-based sense: the adversarial autoencoder learns a multivariate anatomical reference manifold from this midlife cohort and expresses each subject as a deviation from this learned reference, rather than estimating age-conditional normative reference distributions across the lifespan. When applied to an aging target population, the model quantifies individualized deviations from this midlife reference space. The aging population is conceptualized as having originated from a similar healthy middle-aged population but with brain alterations reflecting the cumulative effects of normal aging, latent neuropathologies, and genetic, lifestyle, or environmental risk factors. These multifactorial deviations are summarized using NMF into a reduced set of dominant brain change patterns. The L-map, which captures dynamic brain changes, is derived using a statistical model incorporating longitudinal imaging data, as detailed in subsequent sections. Longitudinal data directly reflect brain changes associated with the active underlying neuropathologic processes at the individual level. By integrating population-based cross-sectional assessments of brain changes with individual trajectories of change over time, CCL-NMF provides a unified decomposition approach with enhanced sensitivity for detecting aging- and disease-related effects.

Normative modeling for estimating the cross-sectional deviation map (C-map)

To estimate cross-sectional deviations (C-map), an adversarial autoencoder (AA)[34,35] was trained on a reference population (S1) and applied it to a target population (S2). Since the AA is optimized to reconstruct only S1-like inputs, the reconstruction error for S2 reflects deviation from the normative reference. This error serves as the basis for the C-map.

The AA architecture comprises an encoder (E) with two hidden layers, each containing 110 neurons and a latent space dimension of 10 neurons. The decoder (D) and the discriminator (D_z) are similarly structured, with two hidden layers of 100 neurons each. The latent space is regularized to match a Gaussian distribution. All hidden layers employ a leaky Rectified Linear Unit with non-linearity, while the latent space and the decoder's output layer utilize a linear activation function.

The AA training has two phases:

1) Reconstruction phase: This phase minimizes the reconstruction loss, ensuring the output closely matches the input. The encoder maps an input feature vector **x** to a latent representation **z**, and the decoder reconstructs it. The reconstruction loss is:

$$L_{\text{recon}} = \left\|\mathbf{x} - \mathrm{D}\left(\mathrm{E}(\mathbf{x})\right)\right\|_2^2 \quad (1)$$

2) Regularization phase: This phase uses adversarial training to constrain the latent representation **z** to follow a prior Gaussian distribution. The discriminator distinguishes samples drawn from the prior Gaussian distribution from latent representations produced by the encoder. The adversarial loss is:

$$L_{\text{adv}} = \mathbb{E}\left[\log\left(\mathrm{D}_{\mathbf{z}}(\tilde{\mathbf{z}})\right)\right] + \mathbb{E}\left[\log\left(1 - \mathrm{D}_{\mathbf{z}}\left(\mathrm{E}(\mathbf{x})\right)\right)\right] \quad (2)$$

where $\tilde{\mathbf{z}}$ denotes samples drawn from the prior Gaussian distribution and $\mathrm{E}(\mathbf{x})$ denotes the latent representation of input **x** produced by the encoder. The encoder minimizes this loss to fool the D_z.

The Adam optimizer is used for 1000 epochs and applies early stopping with 50 epochs of patience. A mini-batch approach is implemented within this gradient descent-based optimizer, with a batch size of 200. A cyclical learning rate enhances the training efficiency, facilitating convergence with fewer epochs. The initial learning rate is 0.0001, with a maximum learning rate of 0.005. The learning rate cycle follows a basic triangular shape with an amplitude decay factor (gamma) of 0.98.

Before the AA, the features were corrected for covariates (sex and ICV) and standardized to z-scores. Both linear correction and z-score models were trained on S1 and were applied to the S2 baseline and longitudinal measures. The S1 was split into three subsets: $S1_{train}$, $S1_{val}$, and $S1_{heldout}$, with a split ratio of 65%, 15%, and 20%, respectively. The AA was trained on $S1_{train}$, validated on $S1_{val,}$ and then applied to $S1_{heldout}$ and S2 baseline measures. The mean squared deviation of an

individual was computed as $\mathrm{MSD} = \frac{1}{R}\sum_{j=1}^{R}\left(x_j - \hat{x}_j\right)^2$, where *R* denotes the number of brain regions, and x_j and $\hat{x}_j$ represent the input and reconstructed values for region *j*, respectively. Individuals with the largest deviations were then selected. Here, since the focus was on atrophy-related volume loss, brain regions with negative deviations (i.e., those with reconstructed values larger than initial values) were not included in the MSD calculation. Individuals in S2 with MSD values exceeding the 75th percentile of MSD in $S1_{heldout}$ were designated as S3, representing those with the greatest deviations from normative space, and were subsequently used in the NMF. The C-map with a size equal to the number of brain regions by the number of S3 individuals includes the deviations. Negative deviations were suppressed by replacing them with 0. The S1, S2, S3 structure is used in both semi-synthetic ($S1_{syn}$, $S2_{syn}$, $S3_{syn}$) and real ($S1_{real}$, $S2_{real}$, $S3_{real}$) data settings.

Statistical method for estimating the longitudinal change map (L-map)

Linear mixed-effects model (LME)[36] is employed to estimate the longitudinal rate of change (L-map) for the S3 individuals with multiple measurements over time. For feature *j*, subject *i*, and timepoint *t*, the model is specified as follows:

$$Y_{i,j,t} = \beta_{0_j} + \beta_{1_j} Time_i + \mathbf{x}_i^{\mathrm{T}} \boldsymbol{\beta}_{2_j} + \gamma_{0_{i,j}} + \gamma_{1_{i,j}} Time_i + \varepsilon_{i,j,t} \quad (3)$$

where $\mathbf{x}_i$ is the covariate vector for subject *i*, β_{0_j} is the population-level intercept for feature *j*, β_{1_j} is the population-level longitudinal slope for feature *j*, and $\boldsymbol{\beta}_{2_j}$ is the vector of fixed-effect coefficients corresponding to the covariate vector $\mathbf{x}_i$. The errors $\varepsilon_{i,j,t}$ are independent and identically distributed with a mean of zero. The subject-specific random intercept and slope parameters, $\gamma_{0_{i,j}}$ and $\gamma_{1_{i,j}}$, are assumed to follow a bivariate normal distribution.

The calculation of each subject's rate of change involves two components: the population-average slope β_{1_j} from the fixed-effects term and the subject-specific random slope $\gamma_{1_{i,j}}$ for subject *i*. So, the final rate of change is given by $\beta_{1_j} + \gamma_{1_{i,j}}$. Time is defined as the interval (in years) between baseline and the last available visit for each subject and is entered into the model as a continuous predictor. By combining fixed and random slopes, the model yields subject-specific rates of change that accommodate variability in follow-up duration and provide stable estimates without requiring balanced schedules.

The resulting L-map is structured with dimensions corresponding to the number of brain regions by the number of individuals with longitudinal measurements, capturing the regional rates of change. Since the focus is on GM atrophy rates, infrequent and low-magnitude positive values were set to zero. To ensure compatibility with the non-negativity constraints required for input data in NMF, the sign of the L-map was inverted, yielding positive values.

Joint NMF

After extracting the C- and L-maps, the joint NMF implementation is carried out. For $\mathbf{X}_C$ of size DxN_C and $\mathbf{X}_L$ of size DxN_L, where D represents the dimensionality of brain features (e.g., ROIs), and N_C (N_L) denotes the number of subjects with cross-sectional deviations (longitudinal change rates), the objective is to extract K components that encapsulate the brain aging patterns using an NMF scheme. This approach operates under the hypothesis that both data types share the same components (dictionary $\mathbf{W}$ of size DxK). The shared dictionary ensures that cross-sectional patterns of brain aging are consistent with the dynamic progression patterns captured by longitudinal measurements. However, cross-sectional and longitudinal measures have distinct loading coefficients: $\mathbf{H}_C$ of size KxN_C for cross-sectional data and $\mathbf{H}_L$ of size KxN_L for longitudinal data. The model can be expressed as:

$$\mathbf{X}_C \approx \mathbf{W}\mathbf{H}_C, \mathbf{X}_L \approx \mathbf{W}\mathbf{H}_L, \text{subject to } \mathbf{W} \geq 0, \mathbf{H}_C \geq 0, \mathbf{H}_L \geq 0 \quad (4)$$

The loss function is:

$$L = \alpha \|\mathbf{X}_C - \mathbf{W}\mathbf{H}_C\|_F^2 + \|\mathbf{X}_L - \mathbf{W}\mathbf{H}_L\|_F^2 \quad (5)$$

where α is a weighting coefficient, determined based on the dataset or application, that regulates the relative influence of C- and L-maps in the dictionary learning process.

This formulation represents a mutually constrained dual factorization of cross-sectional and longitudinal data, optimized using a multiplicative update rule[37,38]:

$$w_{ij} \leftarrow w_{ij} \frac{(\alpha \mathbf{X}_C \mathbf{H}_C^T + \mathbf{X}_L \mathbf{H}_L^T)_{ij}}{(\alpha \mathbf{W}\mathbf{H}_C \mathbf{H}_C^T + \mathbf{W}\mathbf{H}_L \mathbf{H}_L^T)_{ij}} \quad (6)$$

$$h_{ij}^I \leftarrow h_{ij}^I \frac{(\mathbf{W}^T \mathbf{X}_I)_{ij}}{(\mathbf{W}^T \mathbf{W}\mathbf{H}_I)_{ij}}, (I = C, L) \quad (7)$$

Prior to NMF, (non-zero elements of) $\mathbf{X}_C$ and $\mathbf{X}_L$ are rescaled using MinMax scaling to the range [0,1] to ensure uniform feature scaling. The initialization of $\mathbf{W}$, $\mathbf{X}_C$, $\mathbf{X}_L$ matrices is performed with random values drawn from a uniform distribution within [0, 0.5]. The initial matrices are normalized using a diagonal matrix $\mathbf{S}$, derived from the ℓ_2-norms of the columns of $\mathbf{W}_{\text{init}}$:

$$\mathbf{W}'_{\text{init}} = \mathbf{W}_{\text{init}} \mathbf{S}^{-1}, \mathbf{H}'_{C_{\text{init}}} = \mathbf{S}\mathbf{H}_{C_{\text{init}}}, \mathbf{H}'_{L_{\text{init}}} = \mathbf{S}\mathbf{H}_{L_{\text{init}}} \quad (8)$$

This normalization step is repeated at each iteration to ensure numerical stability, promote convergence, and prevent disproportionate scaling of component weights that could lead to component collapse or dominance. An algorithmic pseudocode summarizing the full CCL-NMF pipeline is provided in Supplementary Method 9.

The selection of optimal K is made based on the reproducibility and sparsity indices (Supplementary Method 2). In addition, extended reproducibility analyses were conducted to evaluate the robustness of the decomposition under various sources of variability, including random initialization, bootstrap resampling, and full pipeline reruns with independently trained normative models (Supplementary Method 3).

Out-of-sample estimation of subject-specific loadings

To enable application in external datasets without re-deriving components, which would require a reference cohort, data harmonization, and implementation of the full CCL-NMF pipeline, we fixed the components derived from the discovery dataset and trained separate regression models to easily estimate subject-specific loading coefficients. These models used ROI volumes, age, sex, and ICV as predictors and were fitted independently for each component's loadings. Training and evaluation were performed using five-fold stratified cross-validation based on age, sex, and diagnosis. Model performance was assessed using Spearman correlation between original and approximated loadings. No harmonization was applied to input ROIs to ensure broader applicability. To further validate predictive utility, we applied logistic regression for binary outcomes and Cox proportional hazards models for MCI-to-AD progression and evaluated the models within cross-validation using area under the curve (AUC), concordance index (C-index), and McFadden's and Nagelkerke's pseudo-R^2 metrics. Statistical comparisons between original and regression-based loadings were performed using permutation testing (Supplementary Method 8). By this regression-based approximation, individualized loadings can be efficiently estimated in new subjects, facilitating downstream analyses and clinical translation.

Dataset

Here, CCL-NMF was applied to parse the heterogeneity of aging-related GM changes using 119 predefined GM ROIs (Supplementary Table 1) extracted from baseline T1-w MRI. The T1-w image intensity inhomogeneity was corrected[39], followed by multi-atlas skull-stripping[40], and then the ROIs were segmented using a multi-atlas, multi-warp label fusion-based method[41]. To address outliers and low-quality data, we applied a two-step semi-automated quality control (QC) pipeline: (i) ROI volumes were standardized into z-scores, and cumulative scores were computed to flag potential outliers, and (ii) flagged cases were manually reviewed using MRISnapshot https://github.com/CBICA/MRISnapshot. Scans failing QC criteria were excluded before analysis. All participants gave written informed consent to the study for data acquisition and analyses according to the Declaration of Helsinki. The institutional review board of the University of Pennsylvania approved this project. Sex information was available for all participants and was determined based on self-report as recorded in the original cohort datasets. Regional volumes were adjusted for sex and ICV prior to model estimation, and sex was additionally included as a covariate in downstream statistical analyses to account for potential confounding effects. Participant compensation and study procedures were determined by the original cohort protocols and are described in the respective study documentation.

Semi-synthetic dataset

Semi-synthetic data were generated by simulating atrophy patterns in CN individuals from the UK Biobank (UKBB)[42] cohort to validate the methodology. The seed dataset included 4,517 CN participants (mean age: 51.83±2.33 years, 56.61% females), divided into two subsets: a 20% subset ($S1_{syn}$: *N*=903; mean age: 51.90±2.34 years, 56.59% females) comprised individuals whose data remained unchanged, while the remaining 80% ($S2_{syn}$: *N*=3,614; mean age: 51.81±2.33 years, 56.61% females) underwent simulated atrophy. In $S2_{syn}$, atrophy was simulated

over 40 timepoints (one-year intervals), with varying patterns and onset times drawn from a Gaussian distribution (μ=7, σ=3). Within $S2_{syn}$, 20% of individuals were given synthetic frontal atrophy, 20% occipital, 20% parietal, 20% subcortical, and 20% temporal atrophy. The ROIs affected in each atrophy pattern are listed in Supplementary Table 2. The simulation applied a 1% annual atrophy rate to the ROIs within the pattern and a 0.1% rate to the remaining brain GM ROIs. For the cross-sectional dataset, a single timepoint was randomly selected for each subject from the 40 available, resulting in a synthetic age distribution of 71.30±11.82 years. The L-map of $S3_{syn}$ was derived using an LME model, adjusting for baseline age and ROI volume.

To enhance realism, Gaussian noise was added to each ROI in both map types before NMF (Supplementary Method 10). Three types of NMF experiments were conducted for the five simulated atrophy patterns in the semi-synthetic dataset. First, the NMF was run using solely the C-map. Second, the NMF was run using solely the L-map. Third, the CCL-NMF was run utilizing both C- and L- maps. In the last case, a sensitivity analysis was performed to assess the impact of the α coefficient from Eq. (5) on the accuracy with which the CCL-NMF dictionary captured the simulated atrophy patterns.

To assess the model's ability to identify simulated atrophy patterns, the inner product matrix was computed between the ℓ_2-normalized matrix representing the ground truth simulated atrophy patterns and the ℓ_2-normalized dictionary of the model. The closer the inner product matrix is to the diagonal matrix representing perfect identification/reconstruction of the simulated ground truth atrophy patterns, the better the dictionary captures the simulated atrophy patterns. To quantify the divergence from perfect ground truth reconstruction, the Frobenius norm of the difference between the two matrices (termed divergence matrix) was used; a smaller norm of the divergence matrix indicates a more accurate reconstruction of the ground truth.

iSTAGING dataset

The dataset was drawn from the iSTAGING consortium for the real data analyses. The iSTAGING consortium is a collaborative effort to consolidate neuroimaging, clinical, and cognitive data from more than 70,000 individuals from 24 studies and 38 sites across the adult lifespan. Here, data from the following studies were included: Alzheimer Disease Neuroimaging Initiative (ADNI)[43,44,45,46], Australian Imaging, Biomarker, and Lifestyle (AIBL)[47,48] study, Biomarkers for Older Controls at Risk for Dementia (BIOCARD)[49], Baltimore Longitudinal Study of Aging (BLSA)[50], Coronary Artery Risk Development in Young Adults (CARDIA)[51] study, Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS)[52] study, Open Access Series of Imaging Studies (OASIS)[53], University of Pennsylvania Memory Center cohort (Penn-PMC)[54], Study of Health in Pomerania (SHIP)[55,56], UK Biobank[42], Womens Health Initiative Memory Study (WHIMS)[57,58], and Wisconsin Registry for Alzheimer Prevention (WRAP)[59]. The imaging parameters for each study are presented elsewhere[20]. Inter-site ROI harmonization was performed using the ComBat statistical framework as implemented in the Neuroharmonize toolbox[60] (Supplementary Method 11). Clinical data and cognitive status, where available, were provided by the source study.

The reference population $S1_{real}$ consisted of CN individuals without known CVD-related factors (obesity, hypertension, and diabetes) and age younger or equal to 50 years (*N*=977; mean age 39.88±8.09 years, 54.86% females, 100% CN). The target group $S2_{real}$ comprised individuals older than 50 (*N*=48,949; mean age 65.41±7.92 years, 53.98% females, 94.23% CN). The demographics of $S1_{real}$ and $S2_{real}$ populations by origin study are displayed in Table 1. The L-map was estimated using LME models with the site, baseline age, and ROI volume as covariates. The LME analysis was performed using individuals with three or more longitudinal measures to minimize uncertainty in the rate of change estimation. The joint NMF was run with a varying number of components, $K \in \{2, \ldots, 15\}$.

Statistical analyses

Associations with cognitive, clinical, biomarker, and disease progression measures

Linear and logistic regression analyses were used to evaluate associations between CCL-NMF loading coefficients and AD biomarkers, cognitive performance, the SPARE-AD[61] score, and CVD-related factors. Cox proportional hazards models were employed to assess associations between CCL-NMF loadings and the risk of progression from CN to MCI and from MCI to AD. Hazard ratios (HR) were estimated to quantify the contribution of each component to the risk of cognitive decline. All models included age, sex, and study cohort as covariates, except *APOE4* analyses (age excluded) and ADNI cognitive score analyses (study excluded). Education was additionally included as a covariate in cognitive and disease progression analyses. Bonferroni correction was applied to control for multiple comparisons.

Comparison with Surreal-GAN R-indices

To compare predictive performance between CCL-NMF and Surreal-GAN (Supplementary Method 6), logistic regression was used for binary outcomes (e.g., amyloid, tau, or *APOE4* status) and Cox proportional hazards models were applied to assess MCI-to-AD progression. All models included age, sex, and study as covariates, except for *APOE4* analyses, where age was excluded; education was additionally included in models assessing cognitive progression. Five-fold stratified cross-validation (based on age, sex, and diagnosis) was used to evaluate the models. Predictive performance was assessed using AUC for binary outcomes and C-index for survival analysis. Four model configurations were tested: (1) demographics only, (2) demographics with Surreal-GAN R-indices, (3) demographics with CCL-NMF loadings, and (4) all predictors combined. To further assess explanatory power and formally compare models, we additionally estimated pseudo R^2 values (McFadden's for logistic regression, Nagelkerke's for Cox models) and performed statistical comparisons using permutation tests (testing differences in discrimination and explanatory power) and likelihood ratio tests (formally testing model fit) within the cross-validation framework; full details are provided in Supplementary Method 7.

Data Availability

Raw imaging and clinical data used in this study were provided by several individual studies via data-sharing agreements and are therefore available under restricted access because these agreements do not permit redistribution by the authors. Access to these data can be obtained by applying directly to the respective data providers, as described below. Data from ADNI used in this study are available from the ADNI database (adni.loni.usc.edu) upon registration and compliance with the data usage agreement. Data from the UKBB used in this study are available upon request from the UKBB website (https://www.ukbiobank.ac.uk/). Data from the BLSA study used in this study are available upon request at https://www.blsa.nih.gov/how-apply. Data from the AIBL study used in this study are available upon request at https://aibl.org.au/. Data from the OASIS study used in this study are available upon request at https://www.oasis-brains.org/. Data from the HANDLS study used in this study are available upon request at https://handls.nih.gov/. Data requests for BIOCARD, PENN, WRAP, CARDIA, SHIP, and WHIMS datasets should be directed to the respective study investigators (Marilyn S. Albert, David A. Wolk, Sterling C. Johnson, Lenore J. Launer, Katharina Wittfeld, and Mark A. Espeland). The source data underlying the main figures of this study have been deposited in the Zenodo repository and are available at https://zenodo.org/records/18850194.

Code Availability

Modeling and analyses were conducted in Python (version 3.10). The adversarial autoencoder was implemented in Keras (v2.8) following the architecture described in Pinaya et al.[34]. The linear mixed-effects models were implemented using the statsmodels package (v0.14.5). The joint NMF optimization was implemented using custom code based on the multiplicative update rules from Zhang & Zhang[37] and Lee & Seung[38]. All statistical analyses were performed using statsmodels (v0.14.5), pingouin (v0.5.1), scikit-learn (v1.3.2), and lifelines (v0.25.11) for survival models. Inter-site ROI harmonization was performed with the Neuroharmonize toolbox[60] (Supplementary Method 11), employing the ComBat statistical framework. Finally, component visualization was performed using regional volumetric maps (RAVENS[62]) within the Volume Imaging in Neurological Research, Co-Registration and ROIs included (VINCI64 v5.03[63]) platform.

Out-of-sample estimation of the CCL-NMF loading coefficients is available via the NiChart platform (https://cloud.neuroimagingchart.com/). NiChart enables researchers worldwide to upload their study data, process structural MRI to derive volumetric features (DLMUSE[64] ROIs), and predict CCL-NMF loading coefficients. The code for the pre-trained regression models for loading prediction is available at https://github.com/CBICA/CCL_NMF_Prediction/tree/main. In addition, the full CCL-NMF pipeline—including adversarial autoencoder training, C- and L-map computation, and joint NMF decomposition—together with documentation is publicly available at https://github.com/IoannaSkampardoni/CCL-NMF, complementing the repository already provided for the regression-based estimator.

References

1. Habes, M. *et al.* Advanced brain aging: relationship with epidemiologic and genetic risk factors, and overlap with Alzheimer disease atrophy patterns. *Transl. Psychiatry* **6**, e775 (2016).

2. Eavani, H. *et al.* Heterogeneity of structural and functional imaging patterns of advanced brain aging revealed via machine learning methods. *Neurobiol. Aging* **71**, 41–50 (2018).

3. Zhang, X. *et al.* Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. *Proc. Natl. Acad. Sci.* **113**, E6535–E6544 (2016).

4. Young, A. L. *et al.* Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nat. Commun.* **9**, 4273 (2018).

5. Yang, Z. *et al.* A deep learning framework identifies dimensional representations of Alzheimer's Disease from brain structure. *Nat. Commun.* **12**, (2021).

6. Varol, E., Sotiras, A. & Davatzikos, C. HYDRA: Revealing heterogeneity of imaging and genetic patterns through a multiple max-margin discriminative analysis framework. *Neuroimage* **145**, 346–364 (2017).

7. Dong, A., Honnorat, N., Gaonkar, B. & Davatzikos, C. CHIMERA: Clustering of Heterogeneous Disease Effects via Distribution Matching of Imaging Patterns. *IEEE Trans. Med. Imaging* **35**, 612–621 (2016).

8. Yang, Z. *et al.* Brain aging patterns in a large and diverse cohort of 49,482 individuals. *Nat. Med.* **30**, 3015–3026 (2024).

9. Yang, Z., Wen, J. & Davatzikos, C. Surreal-GAN:Semi-Supervised Representation Learning via GAN for uncovering heterogeneous disease-related imaging patterns. in *International Conference on Learning Representations* (2022).

10. Paatero, P. & Tapper, U. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values†. *Environmetrics* **5**, 111–126 (1994).

11. Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791 (1999).

12. Hassani, A., Iranmanesh, A. & Mansouri, N. Text mining using nonnegative matrix factorization and latent semantic analysis. *Neural Comput. Appl.* **33**, 13745–13766 (2019).

13. Shahnaz, F., Berry, M. W., Pauca, V. P. & Plemmons, R. J. Document clustering using nonnegative matrix factorization. *Inf. Process. Manag.* **42**, 373–386 (2006).

14. Cichocki, A., Zdunek, R. & Amari, S. *New Algorithms for Non-Negative Matrix Factorization in Applications to Blind Source Separation*. *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on* **5**, (2006).

15. Smaragdis, P. & Brown, J. C. Non-negative matrix factorization for polyphonic music transcription. in *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No.03TH8684)* 177–180 (2003). doi:10.1109/ASPAA.2003.1285860

16. Sotiras, A., Resnick, S. M. & Davatzikos, C. Finding imaging patterns of structural covariance via Non-Negative Matrix Factorization. *Neuroimage* **108**, 1–16 (2015).

17. Sotiras, A. *et al.* Patterns of coordinated cortical remodeling during adolescence: associations with functional specialization and evolutionary expansion. *PNAS* **30**, 1–6 (2016).

18. Wen, J. *et al.* Genomic loci influence patterns of structural covariance in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2300842120 (2023).

19. I. Skampardoni, J. Wen, E. Guray, H. Shou, K. N. and C. D. Mutually-Constrained Cross-Sectional and Longitudinal Non-Negative Matrix Factorization: Application to Modeling Brain Aging Trajectories. in *2024 IEEE International Symposium on Biomedical Imaging (ISBI)* 1–5 (2024). doi:10.1109/ISBI56570.2024.10635830

20. Skampardoni, I. *et al.* Genetic and Clinical Correlates of AI-Based Brain Aging Patterns in Cognitively Unimpaired Individuals. *JAMA Psychiatry* **81**, 456–467 (2024).

21. Hwang, G. *et al.* Disentangling Alzheimer's disease neurodegeneration from typical brain aging using MRI and machine learning. *Alzheimer's Dement.* **17**, (2021).

22. Hoyer, P. O. Non-negative Matrix Factorization with Sparseness Constraints. *J. Mach. Learn. Res.* **5**, 1457–1469 (2004).

23. Kuhn, H. W. The Hungarian method for the assignment problem. *Nav. Res. Logist. Q.* **2**, 83–97 (1955).

24. Doshi, J., Erus, G., Habes, M. & Davatzikos, C. DeepMRSeg: A convolutional deep neural network for anatomy and abnormality segmentation on MR images.

25. Tosun, D. *et al.* Identifying individuals with non-Alzheimer's disease co-pathologies: A precision medicine approach to clinical trials in sporadic Alzheimer's disease. *Alzheimers. Dement.* **20**, 421–436 (2024).

26. Hou, B. *et al.* Interpretable deep clustering survival machines for Alzheimer's disease subtype discovery. *Med. Image Anal.* **97**, 103231 (2024).

27. Wen, J. *et al.* Dimensional Neuroimaging Endophenotypes: Neurobiological Representations of Disease Heterogeneity Through Machine Learning. *Biol. Psychiatry* **96**, 564–584 (2024).

28. Feng, X., Provenzano, F. A. & Small, S. A. A deep learning MRI approach outperforms other biomarkers of prodromal Alzheimer's disease. *Alzheimers. Res. Ther.* **14**, 45 (2022).

29. Clerx, L. *et al.* Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment. *Neurobiol. Aging* **34**, 2003–2013 (2013).

30. Woodworth, D. C. *et al.* Dementia is associated with medial temporal atrophy even after accounting for neuropathologies. *Brain Commun.* **4**, fcac052 (2022).

31. Susianti, N. A. *et al.* The impact of medial temporal and parietal atrophy on cognitive function in dementia. *Sci. Rep.* **14**, 5281 (2024).

32. Houben, K. & Jansen, A. When food becomes an obsession: Overweight is related to food-related obsessive-compulsive behavior. *J. Health Psychol.* **24**, 1145–1152 (2019).

33. Maia, T. V, Cooney, R. E. & Peterson, B. S. The neural bases of obsessive-compulsive disorder in children and adults. *Dev. Psychopathol.* **20**, 1251–1283 (2008).

34. Pinaya, W. H. L. *et al.* Using normative modelling to detect disease progression in mild cognitive impairment and Alzheimer's disease in a cross-sectional multi-cohort study. *Sci. Rep.* **11**, 1–13 (2021).

35. Makhzani, A., Shlens, J., Jaitly, N. & Goodfellow, I. J. Adversarial Autoencoders. *ArXiv* **abs/1511.0**, (2015).

36. Liu, X. Chapter 3 - Linear mixed-effects models. in *Methods and Applications of Longitudinal Data Analysis* (ed. Liu, X.) 61–94 (Academic Press, 2016). doi:https://doi.org/10.1016/B978-0-12-801342-7.00003-4

37. Zhang, L. & Zhang, S. A General Joint Matrix Factorization Framework for Data Integration and Its Systematic Algorithmic Exploration. *IEEE Trans. Fuzzy Syst.* **28**, 1971–1983 (2020).

38. Lee, D. & Seung, H. S. Algorithms for Non-negative Matrix Factorization. in *Advances in Neural Information Processing Systems* (eds. Leen, T., Dietterich, T. & Tresp, V.) **13**, (MIT Press, 2000).

39. Tustison, N. J., Cook, P. A. & Gee, J. C. N4Itk: Improved N3 Bias Correction. **29**, 1310–1320 (2011).

40. Doshi, J., Erus, G., Yangming, O., Gaonkar, B. & Davatzikos, C. Multi-Atlas Skull-Stripping. *Acad. Radiol.* **20**, 1566–1576 (2013).

41. Doshi, J. *et al.* MUSE: MUlti-atlas region Segmentation utilizing Ensembles of registration algorithms and parameters, and locally optimal atlas selection. *Neuroimage* **127**, 186–195 (2016).

42. Alfaro-Almagro, F. *et al.* Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* **166**, 400–424 (2018).

43. Kruggel, F., Turner, J. & Muftuler, L. T. Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. *Neuroimage* **49**, 2123–2133 (2010).

44. ADNI | MRI Scanner Protocols. Available at: https://adni.loni.usc.edu/methods/documents/mri-protocols/. (Accessed: 4th January 2023)

45. Brunton, S. *et al.* A voxel-based morphometry comparison of the 3.0T ADNI-1 and ADNI-2 volumetric MRI protocols. *Int. J. Geriatr. Psychiatry* **30**, 531–538 (2015).

46. Gunter, J. L. *et al.* [Ic-P-137]: Adni-3 Mri Protocol. *Alzheimer's Dement.* **13**, 2017 (2017).

47. Rowe, C. C. *et al.* Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol. Aging* **31**, 1275–1283 (2010).

48. Yates, P. A. *et al.* Incidence of cerebral microbleeds in preclinical Alzheimer disease. *Neurology* **82**, 1266–1273 (2014).

49. Albert, M. *et al.* Cognitive changes preceding clinical symptom onset of mild cognitive impairment and relationship to ApoE genotype. *Curr. Alzheimer Res.* **11**, 773–784 (2014).

50. Habes, M. *et al.* White matter lesions: Spatial heterogeneity, links to risk factors, cognition, genetics, and atrophy. *Neurology* **91**, e964–e975 (2018).

51. Lloyd-Jones, D. M. *et al.* The Coronary Artery Risk Development In Young Adults (CARDIA) Study: JACC Focus Seminar 8/8. *J. Am. Coll. Cardiol.* **78**, 260–277 (2021).

52. Waldstein, S. R. *et al.* Differential Associations of Socioeconomic Status With Global Brain Volumes and White Matter Lesions in African American and White Adults: the HANDLS SCAN Study. *Psychosom. Med.* **79**, 327–335 (2017).

53. Marcus, D. S. *et al.* Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *J. Cogn. Neurosci.* **19**, 1498–1507 (2007).

54. Yushkevich, P. A. *et al.* Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. *Hum. Brain Mapp.* **36**, 258–287 (2015).

55. Hosten, N. *et al.* SHIP-MR and radiology: 12 years of whole-body magnetic resonance imaging in a single center. *Healthc.* **10**, 1–35 (2022).

56. Völzke, H. *et al.* Cohort Profile Update: The Study of Health in Pomerania (SHIP). *Int. J. Epidemiol.* **51**, e372–e383 (2022).

57. Resnick, S. M. *et al.* Postmenopausal hormone therapy and regional brain volumes: the WHIMS-MRI Study. *Neurology* **72**, 135–142 (2009).

58. Kuller, L. H. *et al.* Relationship of hypertension, blood pressure, and blood pressure control with white matter abnormalities in the Women's Health Initiative Memory Study (WHIMS) - MRI trial. *J. Clin. Hypertens.* **12**, 203–212 (2010).

59. Birdsill, A. C. *et al.* Regional white matter hyperintensities: aging, Alzheimer's disease risk, and cognitive function. *Neurobiol. Aging* **35**, 769–776 (2014).

60. Pomponio, R. *et al.* Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *Neuroimage* **208**, 116450 (2020).

61. Davatzikos, C., Xu, F., An, Y., Fan, Y. & Resnick, S. M. Longitudinal progression of Alzheimers-like patterns of atrophy in normal older adults: The SPARE-AD index. *Brain* **132**, 2026–2035 (2009).

62. Davatzikos, C., Genc, A., Xu, D. & Resnick, S. M. Voxel-based morphometry using the RAVENS maps: Methods and validation using simulated longitudinal atrophy. *Neuroimage* **14**, 1361–1369 (2001).

63. S. Vollmar, J. Cizek, M. Sué, J. Klein, A. H. Jacobs, K. H. VINCI -Volume Imaging in Neurological Research, Co-Registration and ROIs included. in *Forschung und wissenschaftliches Rechnen 2003 (Kremer K, Macho V, eds)* 115–131 (2004).

64. Bashyam, V. et al. Robust Brain Segmentation in Seconds using Deep Learning: Validation Across 71,391 Brain MRI Scans from 14 Neuroimaging Studies. *Radiol. Artif. Intell.* **In Review**,

Acknowledgments

The iSTAGING study is a multi-institutional effort funded by the National Institute on Aging (NIA) by RF1 AG054409 (C. Davatzikos). The BLSA neuroimaging study is funded by the Intramural Research Program, NIA, National Institutes of Health (NIH), and by HHSN271201600059C. K.A.W., M.R.D., and S.M.R. were supported by the Intramural Research Program of the NIH, NIA. The BIOCARD study is in part supported by NIH grant U19-AG033655. SHIP is part of the Community Medicine Research net of the University of Greifswald, which is funded by the Federal Ministry of Education and Research (grant nos. 01ZZ9603, 01ZZ0103 and 01ZZ0403), the Ministry of Cultural Affairs and the Social Ministry of the Federal State of Mecklenburg-West Pomerania. MRI scans in SHIP-START and SHIP-TREND have been supported by a joint grant from Siemens Healthineers and the Federal State of Mecklenburg-West Pomerania. The Women's Health Initiative was funded by the National Heart, Lung and Blood Institute of the NIH, US Department of Health and Human Services. Contracts HHSN268200464221C and N01-WH-4–4221 provided additional support. The WHIMS was funded in part by Wyeth Pharmaceuticals. The HANDLS study is supported by the NIA Intramural Research Program, NIH, Project ZIA-AG000513. The HANDLS Scan substudy is supported by NIH grants 1RO1AG034161, 2P30AG028747–14S1 and 1R56AG064088–01A1. This research has been conducted using the UKBB Resource under application no. 35148. Data used in this article were in part obtained from the ADNI database (adni.loni.usc.edu). A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.
ADNI is funded by the NIA, 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; Biogen; Bristol-Myers Squibb; CereSpir; Cogstate; Eisai; Elan Pharmaceuticals; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche and its affiliated company Genentech; Fujirebio; GE Healthcare; IXICO; Janssen Alzheimer Immunotherapy Research & Development; Johnson & Johnson Pharmaceutical Research & Development; Lumosity; Lundbeck; Merck & Co; Meso Scale Diagnostics; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. I.S. had full access to all the data in the study and is responsible for the integrity of the data and the accuracy of the data analysis. Funding sources had no role in the study design, data collection, analysis, interpretation, or writing of the study report.

Author Contributions Statement

C.D. conceived the project and supervised the research. I.S. co-designed the project, developed the computational methods, performed the data analysis, and drafted the manuscript. A.G. and

K.B. integrated the model into NiChart software. G.E., I.M.N., Z.Y., M.R.D., K.A.W., R.M., S.T.G., S.M.R., H.S., and K.N. provided critical revisions. C.D., I.S., G.E., I.M.N., Z.Y., M.R.D., K.A.W., A.G., K.B., R.M., S.T.G., S.M.R., H.S., and K.N. reviewed and approved the final version of the manuscript.

Competing Interests Statement

I.M.N. has received fees for educational speaking from Biogen and for an advisory board for Eisai. K.A.W. is an Associate Editor at Alzheimer's & Dementia, a member of the Editorial Board of Annals of Clinical and Translational Neurology, and on the Board of Directors of the National Academy of Neuropsychology. K.A.W. has given unpaid talks on behalf of SomaLogic. These activities are outside the scope of the present work. All other authors declare no competing interests.

Tables

Table 1: Demographic summary of reference ($S1_{real}$) and target ($S2_{real}$) population. For age, mean and range (min–max) are presented. Abbreviation: CN, Cognitively Normal.

	Sample size		Age (years)		Sex (%males)		Diagnosis (%CN)	
	Target	Reference	Target	Reference	Target	Reference	Target	Reference
ADNI	2391	-	73.1 (54.27 - 91.31)	-	52.4	-	36.4	-
AIBL	922	4	73.1 (54.25 -96)	45.45 (43.16-49.07)	43.5	25	76	100
BIOCARD	259	-	60.8 (50.06 - 86.27)	-	40.9	-	97.7	-
BLSA	916	100	70.1 (51-93)	40.54 (22-50)	47.3	42	97.5	100
CARDIA	534	170	53.9 (51-61)	47.22 (43-50)	46.4	50.6	100	100
HANDLS	147	33	58.6 (50.1-72.4)	42.89 (35.9-49.9)	44.9	57.6	100	100
OASIS	1097	10	71.8 (50.08 - 101.79)	47 (42.66-49.24)	45	20	73.3	100
PENN	959	-	73.9 (51-95)	-	43.1	-	20.8	-
SHIP	1810	660	63 (51-90)	37.6 (21-50)	48	44.1	100	100
UK BIOBANK	38582	-	64.5 (50-82.27)	-	47.1	-	100	-
WHIMS	1080	-	69.6 (64-79)	-	0	-	100	-
WRAP	252	-	62.1 (50.3-77.7)	-	29	-	99.6	-

Figure Legends/Captions

Figure 1: Overview of the CCL-NMF framework and its application to brain atrophy. A–C) Coupled Cross-sectional and Longitudinal Non-negative Matrix Factorization (CCL-NMF) integrates baseline and longitudinal MRI data from an older target population. Individuals with baseline and longitudinal scans are used to construct two input matrices: the C-map ($\mathbf{X}_C$) representing cross-sectional deviations from a normative reference population, and the L-map ($\mathbf{X}_L$) capturing individualized gray matter atrophy rates over time. These maps are jointly decomposed into a shared dictionary of spatial atrophy components (**W**) and into separate loading matrices $\mathbf{H}_C$ and $\mathbf{H}_L$, which quantify subject-level expression of each component in the cross-sectional and longitudinal domains, respectively. D) The resulting components capture distinct spatial patterns of gray matter atrophy, and the subject-specific loadings demonstrate biological relevance, support reliable out-of-sample estimation via the NiChart platform, and improve prediction of biomarker and disease progression outcomes compared to models based solely on cross-sectional data. Although this figure illustrates gray matter atrophy, the framework is general and can be extended to other monotonic brain changes, such as lesion burden or molecular pathology accumulation.

Figure 2: Results in semi-synthetic data. From left to right, matrices show the inner product matrix between the ℓ_2-normalized matrix of the ground truth simulated atrophy patterns and the ℓ_2-normalized dictionaries estimated by C-NMF (C-map only), L-NMF (L-map only), and CCL-NMF (jointly using C- and L-maps). Brighter (orange/white) cells indicate higher inner product (greater similarity), and darker (purple/black) cells indicate lower similarity. A pronounced diagonal structure indicates accurate recovery of the simulated brain atrophy patterns.

Figure 3: Determination of the number of CCL-NMF components. (A) Reproducibility index as a function of the number of components (*K*). The dataset was divided into two subsets with comparable age and sex distributions, and CCL-NMF was applied independently to each subset. Components from the two decompositions were matched using the Hungarian algorithm, and their similarity was quantified using the inner product (reproducibility index; range 0–1). Points represent the mean reproducibility index across matched component pairs, and error bars indicate ± standard deviation (SD) across matched pairs. (B) CCL-NMF dictionary sparsity as a function of the number of components. Higher values indicate more spatially sparse component patterns. Additional methodological details are provided in Supplementary Method 2.

Figure 4: Spatial atrophy patterns captured by CCL-NMF components. Each of the seven components reflects a distinct spatial pattern of brain atrophy. The color bar indicates the relative contribution weight of each region to the component. Warmer (cooler) colors indicate higher (lower) regional contributions.

Figure 5: Volumetric longitudinal trajectories reveal accelerated atrophy in individuals with high longitudinal CCL-NMF loadings, highlighting the model's capacity to track neurodegeneration. Representative brain regions include (A) hippocampus (CCL-NMF2), (B) precuneus (CCL-NMF4), and (C) insular cortex (CCL-NMF5). Data points represent individual longitudinal measurements of regional brain volume plotted against age. Red circular markers denote subjects with high longitudinal component loadings (top 10th percentile), whereas blue square markers denote subjects with low longitudinal loadings (bottom 10th percentile). Subjects were randomly sampled from these percentile groups for visualization. Linear regression lines illustrate volume trajectories over time, with steeper declines observed among subjects with higher longitudinal loadings, thereby revealing the ability of CCL-NMF to capture progressive neurodegenerative changes. ROI volumes are reported in mm^3 and age in years.

Figure 6: Associations between cross-sectional CCL-NMF loadings and clinical, cognitive, and vascular measures. Cross-sectional CCL-NMF loadings were tested for association with (A) AD-related measures, (B) age, (C) cognitive performance, (D) risk of cognitive impairment progression, and (E) cardiovascular risk factors. For forest plots, points indicate estimated effect sizes (odds ratios (OR) or hazard ratios (HR)) and horizontal lines represent the corresponding 95% confidence intervals (CI); values are displayed on a log scale where indicated. For bar plots, bars represent the estimated regression coefficients (β) from linear models. Statistical significance was evaluated using two-sided tests (Wald tests for logistic and Cox regression; t-tests for linear regression). Statistical significance is indicated as $p < 0.0029$ (*) and $p < 5.9\times10^{-6}$ (**), after Bonferroni correction for 17 tests. *N* denotes the number of independent participants included in each analysis. For progression analyses (Panel D), Cox proportional hazards models were used, and *N* represents the total number of individuals included in the model (including both progressors and non-progressors). All models included age, sex, and study cohort as covariates, except *APOE* ε4 analyses (age excluded) and ADNI cognitive score analyses (study excluded). Education was additionally included as a covariate in cognitive (Panel C) and disease progression analyses (Panel D). The average follow-up time was 5.37 ± 4.29 years for CN→MCI conversion and 2.6 ± 2.65 years for MCI→AD progression. Definitions for AD biomarkers and CVD-related factors are provided in Supplementary Method 5. Abbreviations: SPARE-AD, Spatial Pattern of Abnormality for Recognition of Early Alzheimer's Disease; WMH, White Matter Hyperintensities; CN, Cognitively Normal; MCI, Mild Cognitive Impairment; ADNI, Alzheimer's Disease Neuroimaging Initiative; ADNI-MEM, ADNI memory composite; ADNI-VS, ADNI visuospatial functioning composite; ADNI-LAN, ADNI language composite; ADNI-EF, ADNI executive function composite; ADAS-COG, Alzheimer's Disease Assessment Scale-Cognitive Subscale; *APOE*, Apolipoprotein E.

Figure 7: Prediction of AD markers, cardiovascular disease risk, and cognitive impairment progression using imaging-derived components. Models incorporating CCL-NMF loadings outperform demographics-only models and those using Surreal-GAN R-indices, demonstrating the added predictive value of integrating longitudinal information. Bars represent mean predictive performance (area under the curve (AUC) or concordance index (C-

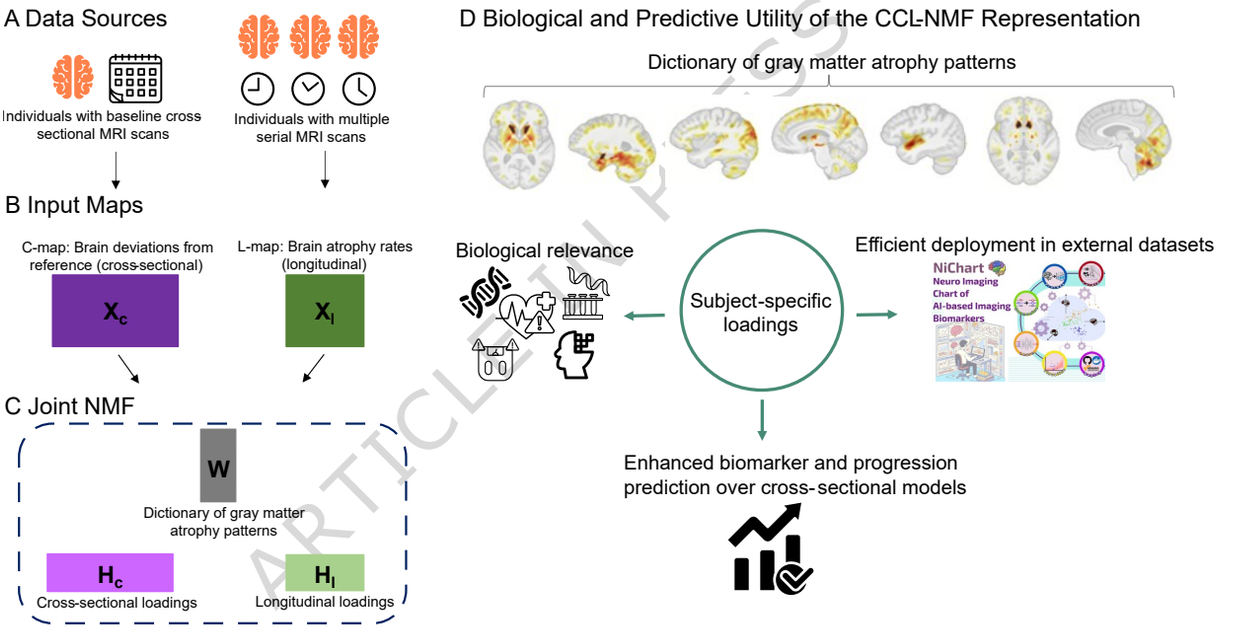
index)) across cross-validation folds, and error bars indicate ± standard deviation (SD) across folds. Outcomes include *APOE* ε4 status, amyloid positivity, tau positivity, obesity, hypertension, and MCI→AD progression. Four model configurations are shown: (1) demographics only, (2) demographics and R-indices, (3) demographics and CCL-NMF loadings, and (4) a full model including all predictors. All models included age, sex, and study cohort as covariates, except for *APOE* ε4 status (age excluded); education was additionally included as a covariate in the MCI→AD progression model. *N* indicates the number of independent participants included in each prediction analysis, as shown below each outcome. Definitions of biomarkers and cardiovascular risk factors are provided in Supplementary Method 5. Abbreviations: MCI, Mild Cognitive Impairment; *APOE*, Apolipoprotein E.

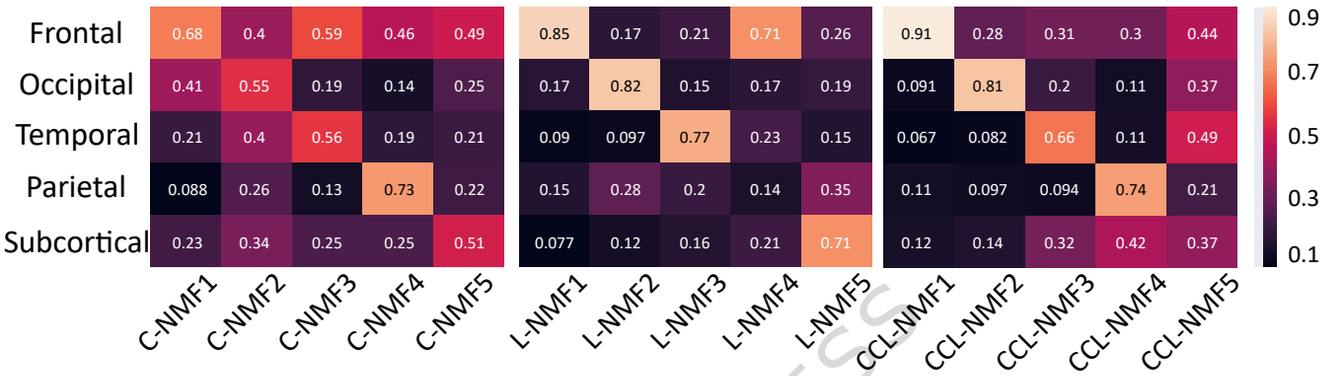
Figure 8: Correspondence between original and approximated CCL-NMF loading coefficients. (A) Spearman correlations between original and regression-based approximated cross-sectional loading coefficients across subjects in the cross-sectional dataset (*N*=13,950). (B) Spearman correlations between original and approximated longitudinal loading coefficients across subjects with longitudinal data (*N*=1,063). Each cell represents the correlation between an original component loading (rows; upper-case labels) and its regression-based approximation (columns; lower-case labels). The high correlations (0.8–0.93 for cross-sectional and 0.9–0.97 for longitudinal loadings) indicate that the approximated loadings accurately preserve the original components' structure, ensuring their reliability for out-of-sample applications.

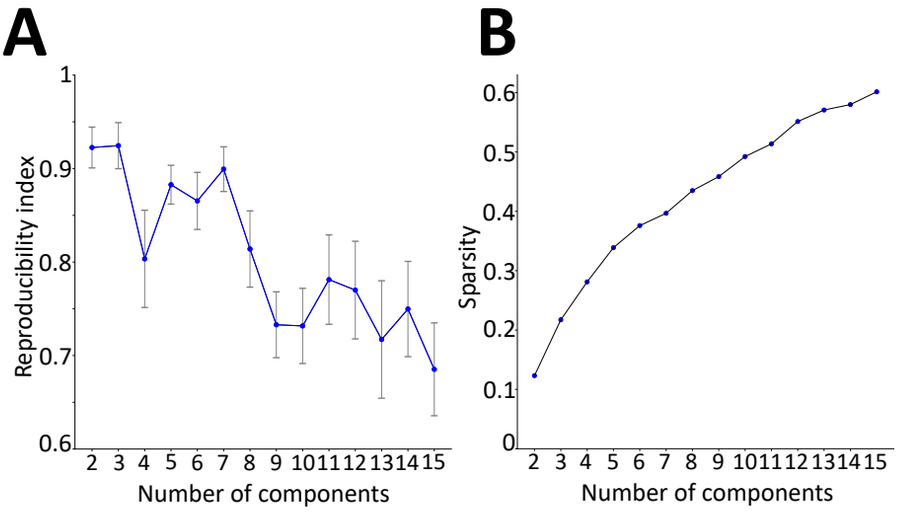
Figure 9: Schematic representation of the CCL-NMF framework, including C-map and L-map estimation and joint decomposition. An adversarial autoencoder (AA) is used to estimate the cross-sectional deviation map (C-map) of the target population (here, an aging cohort) relative to a normative reference population (here, a healthy middle-aged cohort). Linear mixed-effects (LME) models are applied to individuals with longitudinal measurements to estimate the rate of change map (L-map). The resulting C-map ($\mathbf{X}_C$) and L-map ($\mathbf{X}_L$) are jointly decomposed via non-negative matrix factorization (NMF) into a shared dictionary ($\mathbf{W}$) that captures distinct patterns of brain change and loading matrices ($\mathbf{H}_C$ and $\mathbf{H}_L$) that reflect the individual expression levels of each component. This coupled factorization is defined as $\mathbf{X}_C \simeq \mathbf{W}\mathbf{H}_C$ and $\mathbf{X}_L \simeq \mathbf{W}\mathbf{H}_L$ with $\mathbf{W} \geq 0$, $\mathbf{H}_C \geq 0$, and $\mathbf{H}_L \geq 0$. The shared dictionary links cross-sectional deviations and longitudinal change within a common component space. This schematic complements the conceptual overview in Fig. 1 by detailing the methodological steps involved in estimating input maps and performing the coupled decomposition. The symbol ‘x’ indicates matrix multiplication.

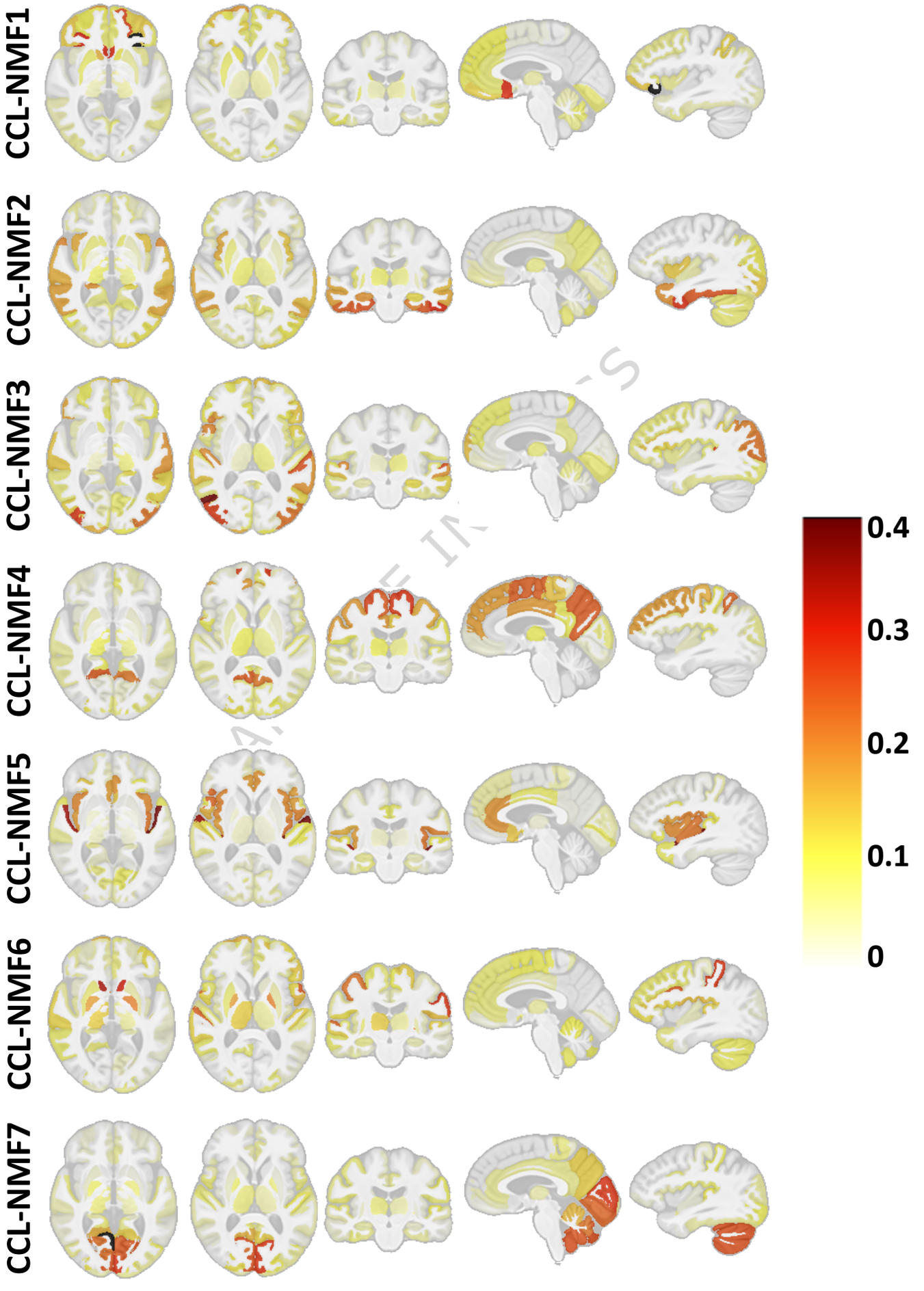
Editor's Summary: A machine-learning framework integrating cross-sectional and longitudinal brain imaging reveals distinct brain aging trajectories in ~49,000 individuals and links them to Alzheimer's disease biomarkers, cognition, and cardiovascular risk factors.

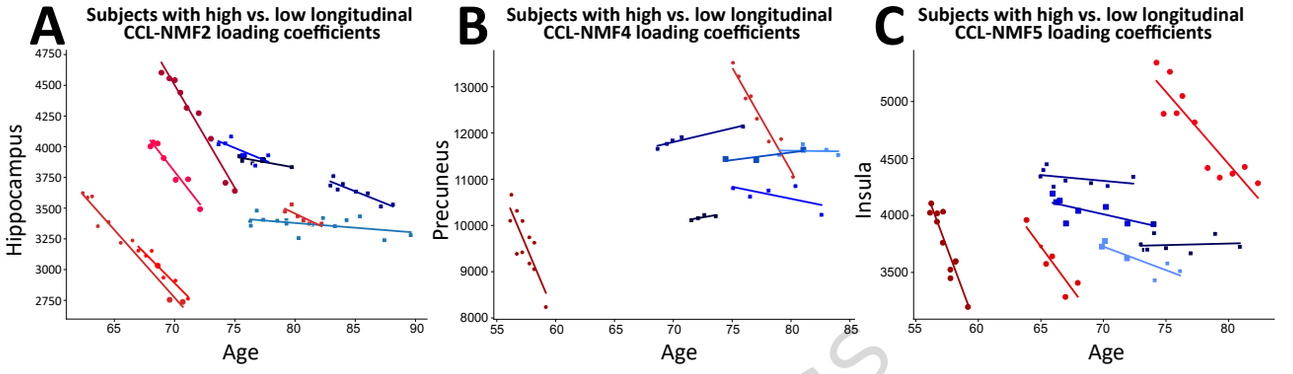
Peer review information: *Nature Communications* thanks Jia Guo, Tobias Kaufmann, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

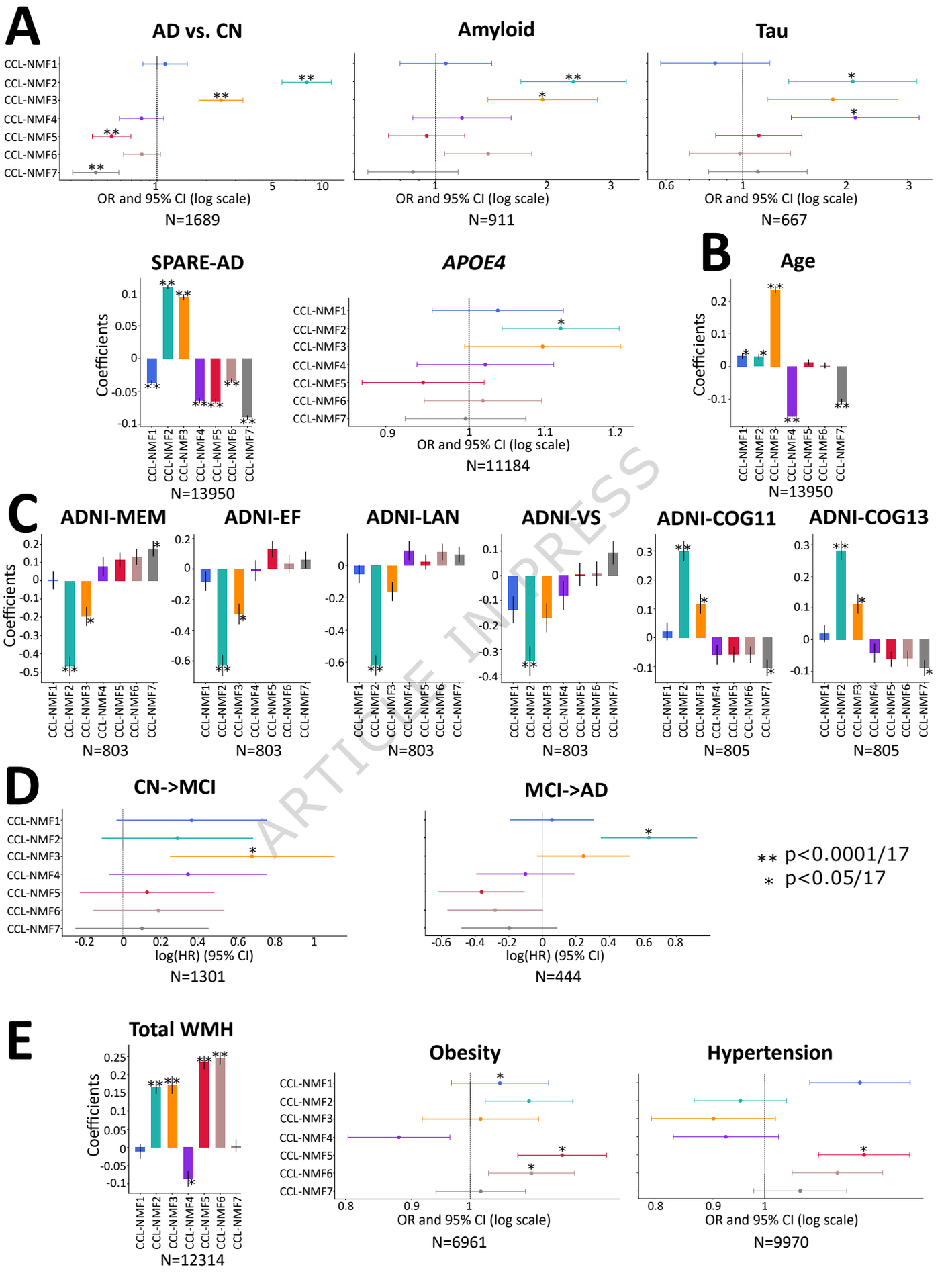


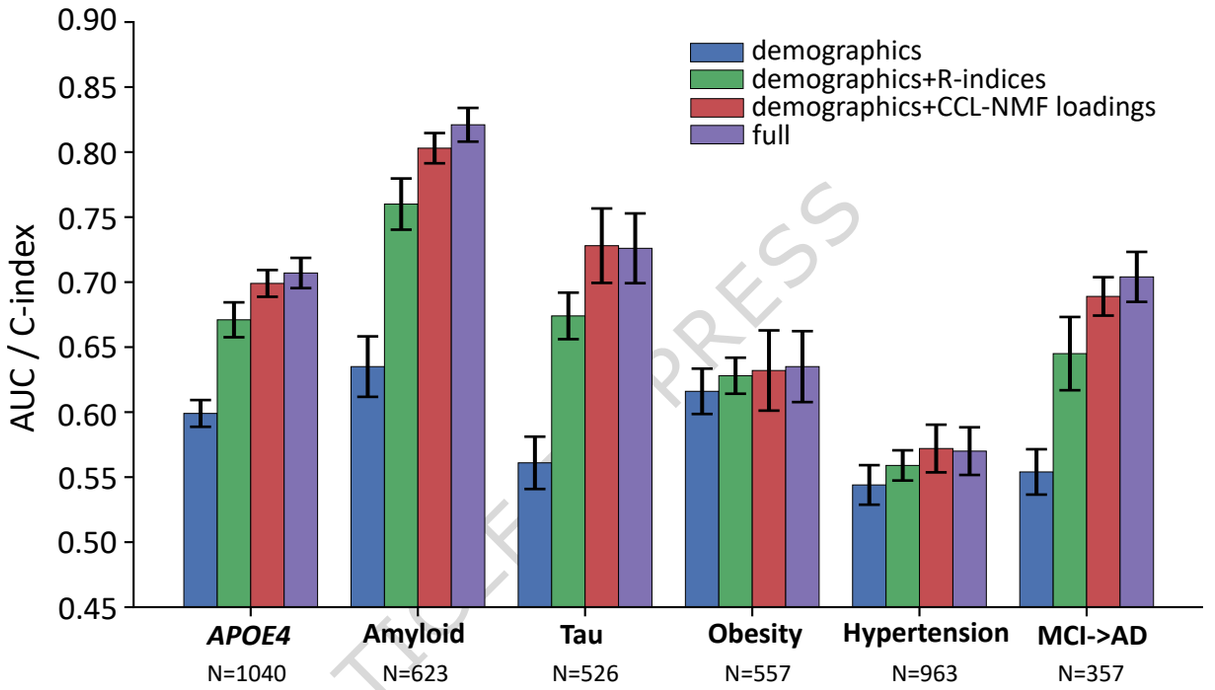


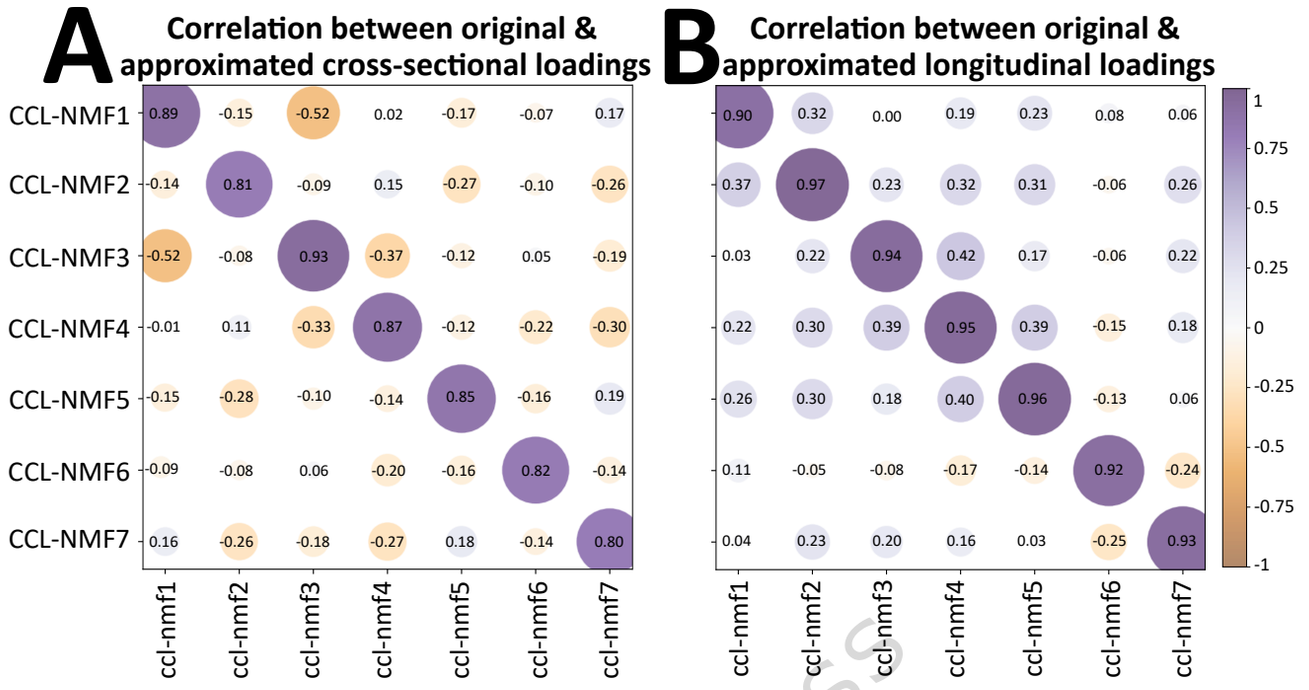


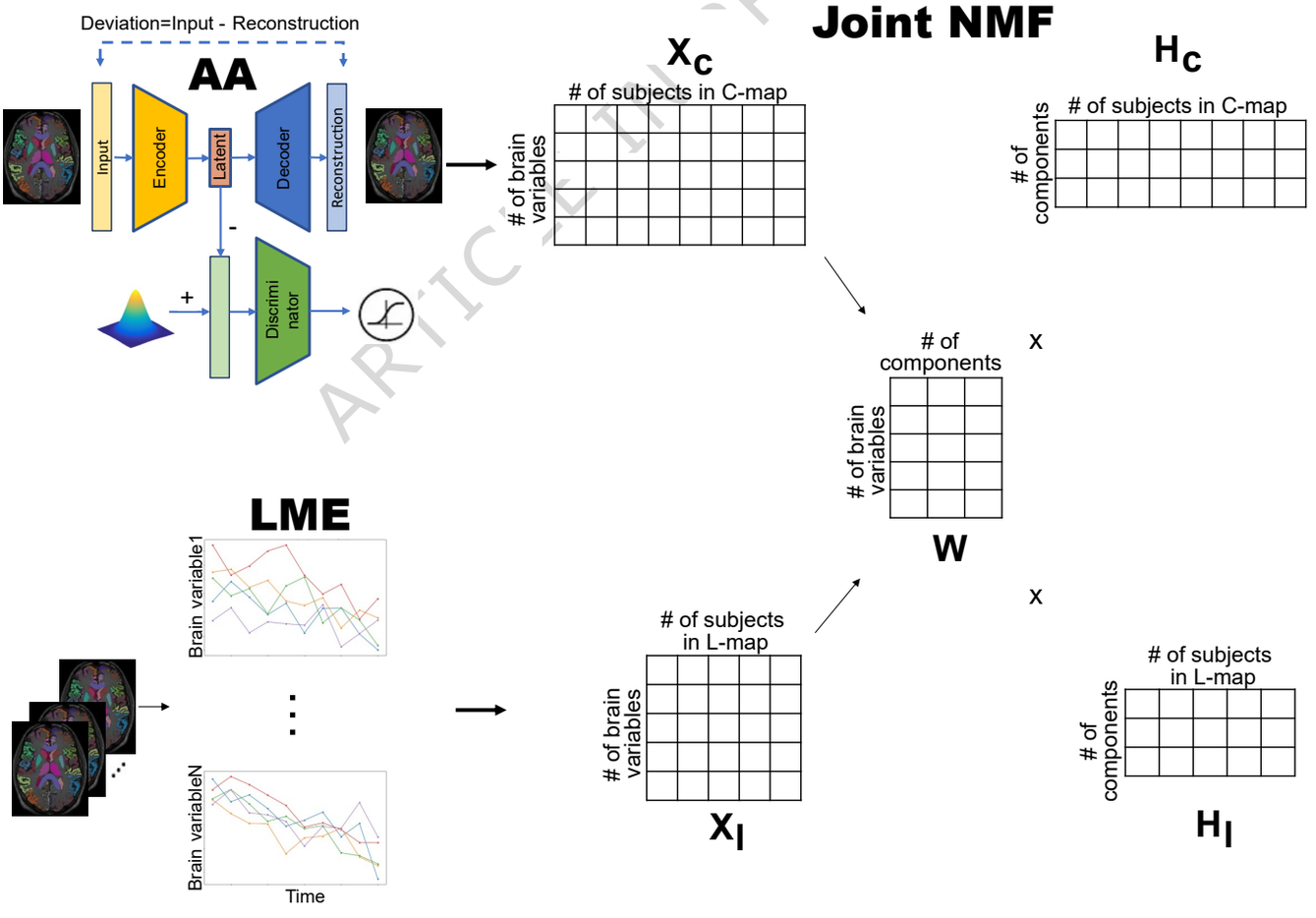

CCL-NMF1
CCL-NMF2
CCL-NMF3
CCL-NMF4
CCL-NMF5
CCL-NMF6
CCL-NMF7
0.4
0.3
0.2
0.1
0










Deviation=Input - Reconstruction
AA
Input
Encoder
Latent
Decoder
Reconstruction
-
+
Discrimi
nator
Joint NMF
X_C
# of subjects in C-map
# of brain variables
H_C
# of components
# of subjects in C-map
x
# of components
# of brain variables
W
x
LME
Brain variable1
Brain variableN
Time
# of subjects in L-map
# of brain variables
X_I
# of subjects in L-map
# of components
H_I