Spectral normative modeling of brain structure

Sina Mansour L. [1,2,3,4,5,6] ✉, **Maria A. Di Biase** [2,7,8], **Chen Zhang** [1], **Fang Tian** [1,3,4,5,6,9,10], **Shaoshi Zhang** [1,3,4,5,6,9,10,11,12], **Hongwei Yan** [1], **Aihuiping Xue** [1], **Joanna Su Xian Chong** [1,3,4,5], **Niousha Dehestani** [1,3,4,5], **Eric Kwun-Kei Ng** [1,3,4,5], **Fang Ji** [1,3,4,5], **Xing Qian** [1,3,4,5], **Yichi Zhang** [1,3,4,5], **Wen Liang Loh** [1,3,4,5], **Joice Sin Yi Tham** [1,3,4,5], **Voon Hao Lew** [1,3,4,5], **Samantha Hui Fang Neo** [1,3,4,5], **Fiona Jia Wen Goh** [1,3,4,5], **Narayanaswamy Venketasubramanian** [13,14], **Eddie Chong** [14], **Nagaendran Kandiah** [15], **Ai Peng Tan** [16,17,18,19], **Michael J. Meaney** [20], **Marielle V. Fortier** [18], **Yap Seng Chong** [18,21], **Woon-Puay Koh** [3], **Vanessa Cropley** [22,23], **Jakob Seidlitz** [24,25,26,27], **Aaron Alexander-Bloch** [24,25,26,27], **Richard Bethlehem** [28], **Christopher Chen** [14,29], **Juan Helen Zhou** [1,3,4,5,9,11], **the Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing** [*], **the Alzheimer's Disease Neuroimaging Initiative** [**], **the Lifespan Brain Chart Consortium**, **B.T. Thomas Yeo** [1,3,4,5,6,9,10,11,30] ⚖, ✉, and **Andrew Zalesky** [2,31,] ⚖, ✉

[1] *Centre for Sleep & Cognition & Centre for Translational Magnetic Resonance Research, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[2] *Systems Neuroscience Lab, Department of Psychiatry, The University of Melbourne, Parkville, Victoria, Australia*
[3] *Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[4] *Human Potential Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[5] *Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[6] *Institute for Digital Medicine (WisDM), Yong Loo Lin School of Medicine, National University of Singapore, Singapore*
[7] *Stem Cell Disease Modelling Lab, Department of Anatomy and Physiology, The University of Melbourne, Parkville, Victoria, Australia*
[8] *Psychiatry Neuroimaging Laboratory, Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, USA*
[9] *Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore*
[10] *N.1 Institute for Health, National University of Singapore, Singapore, Singapore*
[11] *Integrative Sciences and Engineering Programme (ISEP), National University of Singapore, Singapore, Singapore*
[12] *Center for Brain Circuit Therapeutics, Brigham and Women's Hospital, Harvard Medical School, Boston, USA*
[13] *Raffles Neuroscience Centre, Raffles Hospital, Singapore, Singapore*
[14] *Memory Aging and Cognition Centre, National University Health System, Singapore, Singapore*
[15] *Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore*
[16] *Department of Diagnostic Imaging, National University Health System Singapore, Singapore, Singapore*
[17] *Department of Diagnostic Radiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[18] *Institute for Human Development and Potential, Agency for Science, Technology and Research in Singapore, Singapore, Singapore*
[19] *Yeo Boon Khim Mind Science Centre, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[20] *Douglas Mental Health University Institute, McGill University, Montreal, Canada*
[21] *Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[22] *Centre for Youth Mental Health, University of Melbourne, Parkville, Victoria, Australia*
[23] *Orygen Youth Health, Orygen, Parkville, Victoria, Australia*
[24] *Lifespan Brain Institute, Children's Hospital of Philadelphia, Philadelphia, PA United States*
[25] *Department of Child and Adolescent Psychiatry and Behavioral Science, Children's Hospital of Philadelphia, Philadelphia, PA United States*
[26] *Department of Psychiatry, University of Pennsylvania, Philadelphia, PA United States*
[27] *Brain-Gene Development Laboratory, Lifespan Brain Institute at Children's Hospital of Philadelphia and University of Pennsylvania, Philadelphia, PA, United States*
[28] *Department of Psychology, University of Cambridge, Cambridge, United Kingdom*
[29] *Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[30] *Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States*
[31] *Department of Biomedical Engineering, The University of Melbourne, Parkville, Victoria, Australia*
[*] *Data used in this article were obtained in part from the Australian Imaging Biomarkers and Lifestyle flagship study of ageing (AIBL). See* https://aibl.org.au/ *and the Acknowledgments for further details.*
[**] *Data used in this article were obtained in part from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (*adni.loni.usc.edu*). Additional details are provided in the Acknowledgments.*
⚖ *These authors contributed equally to this work.*
✉ *E-mail: sina.mansour.lakouraj@gmail.com; thomas.yeo@nus.edu.sg; azalesky@unimelb.edu.au*

Abstract: Normative modeling in neuroscience aims to characterize interindividual variation in brain phenotypes and establish reference ranges, or brain charts, against which individuals can be compared. Normative models are typically limited to coarse spatial scales due to computational constraints, limiting their spatial specificity. Furthermore, dependence on fixed parcellation atlases limits their adaptability to alternative parcellation schemes. To overcome these key limitations, we propose *spectral normative modeling* (SNM), which leverages brain eigenmodes to efficiently generate normative ranges for arbitrarily defined regions of interest. Training SNM on over 78,000 healthy brain scans, we generate accurate lifespan thickness growth charts across different spatial scales, from millimeters to the whole brain. These charts reveal three principal thickness growth gradients, aligning neurotypical cortical change with established anatomical, genetic, and functional hierarchies. We further demonstrate SNM's utility by elucidating high-resolution individual cortical atrophy patterns that characterize the heterogeneous expression of neurodegeneration in Alzheimer's disease. SNM lays the groundwork for a new generation of spatially precise brain charts, offering substantial potential to drive advances in individualized precision medicine.

Keywords: Normative Modeling; Brain Charts; Graph Signal Processing; Brain Eigenmodes

1 Introduction

Normative modeling aims to estimate population-level reference ranges for variation in a phenotype of interest, providing a methodological framework to quantify individual deviations from expected patterns[1–3]. Individuals can be benchmarked against established normative ranges while accounting for relevant covariates, such as age, sex, or other demographic, clinical, or biological variables, to determine whether, and to what extent, they fall outside a typical healthy range[4,5]. Studying such individualized deviations in relation to population trajectories has become an important goal in neuroscience, motivating the development of normative models for numerous whole-brain and regional phenotypes, such as cortical thickness and volume[6]. Prior research has demonstrated that these models can capture the heterogeneous patterns of individual deviation in brain structure[4,6]. Consequently, normative modeling has emerged as a powerful approach with significant potential to advance precision medicine[7–14].

Normative brain charts for magnetic resonance imaging (MRI) phenotypes can, in principle, be established at the highly detailed spatial resolution at which the MRI scan is acquired. This would enable highly localized and spatially specific inference about an individual's deviations in a cortical phenotype. However, in practice, considerations related to computational cost, high-resolution phenotype reliability, and model scalability have motivated most normative modeling studies to operate at coarser spatial resolutions. Conventional normative approaches are designed to estimate ranges for a single phenotypic summary statistic, such as mean cortical thickness averaged across the whole cortex[2,15,16]. To reach higher spatial specificity, normative brain charting studies typically repeat model-fitting for several regional summary statistics defined over a predetermined brain atlas[6,17].

Alternatively, a mass-univariate application of conventional techniques over all voxels/vertices has been used to produce norms at high spatial resolution[8,13,18]. While enabling fine-grained inference, these approaches are sensitive to voxel-level noise and incur substantial computational cost, especially in population-wide studies using large-scale imaging biobanks. Moreover, with current normative modeling approaches, determining a reference range for a new spatial region of interest requires fitting a separate normative model anew, a process often hindered by limited access to the original training data that is typically unavailable to end-users. Establishing a methodological framework that efficiently alleviates such limitations will enable a more principled and efficient mapping of brain charts at high spatial resolution compared to the current mass-univariate approach.

Developing high-resolution normative models is challenging, particularly due to the high dimensionality of the feature space, i.e. hundreds of thousands of vertices on the cortical surface mesh. Importantly, these challenges are not unique to normative modeling, but are shared broadly across neuroimaging analyses that aim to balance spatial precision with computational tractability. For instance, spatial dependencies across vertices/voxels undermine independence assumptions and complicate the development of models that accurately explain high-resolution statistical interdependencies[19–21]. As such, finding an appropriate low-dimensional encoding of high-resolution cortical information may enable the development of computationally tractable techniques for high-resolution normative models.

Through recent advances in brain signal processing, eigenmodes constructed from the brain's geometry and connectivity have yielded promising basis functions that can summarize phenotypic variations on the cortical surface[22–24]. As such, eigenmodes can provide a solution to high-resolution normative modeling of brain phenotypes. Spatial variation in a cortical phenotype can be captured using a lower dimensional graph spectral embedding[25,26]. We exploit this parsimonious eigenmode basis to establish normative models on the coefficients of cortical phenotypes expressed in this lower dimensional latent space. By formulating a method that relates the normative range of an arbitrary region of interest to eigenmode normative ranges, we develop a computationally efficient method that simultaneously estimates normative ranges over multiple spatial granularities.

We introduce spectral normative modeling (SNM), a framework for constructing normative reference brain charts that are independent of spatial resolution or parcellation. SNM represents high-dimensional brain phenotypes, such as cortical thickness, using a compact eigenmode representation, enabling reconstruction of normative ranges at arbitrary spatial scales. We validate SNM's performance across multiple spatial scales and apply it to model cortical thickness norms in over 78,000 healthy individuals. Using this model, we identify three principal gradients of lifespan thickness growth that link to previously known brain hierarchies. Finally, we demonstrate the utility of this framework in characterizing individual cortical thickness deviations associated with cognitive impairment in Alzheimer's disease, highlighting its ability to generalize to unseen samples and elucidate individual heterogeneity in cortical atrophy.

2 Results

2.1 SNM: An Efficient Framework for Multi-Scale Normative Modeling

The fundamental characteristics of many engineered and natural systems can be modeled by their structural eigenmodes, which offer a simplified yet powerful means of capturing the system's behavior. Recent advances in neuroscience have similarly demonstrated that brain eigenmodes provide a parsimonious basis for characterizing cortical information[22,25,27]. By leveraging such basis sets, we aim to enhance both the efficiency and spatial versatility of conventional normative models. In particular, we use brain connectivity eigenmodes, which naturally extend across both cortical hemispheres while preserving key neuroanatomical landmarks, such as homotopic symmetries and the separation of cortical lobes. This is especially advantageous given that pathological brain alterations often propagate along the brain's structural network[28–34], making connectivity eigenmodes well suited to capture normative deviations induced by mechanisms of axonal propagation. However, it is important to note that the methodological advances introduced in this work are not limited to connectivity eigenmodes and can be readily generalized to any basis set for information reconstruction.

To establish robust normative thickness trajectories across the human lifespan, we collated preprocessed structural MRI data from 78,405 healthy individuals (5–95 years) across 30 datasets and 189 sites (Fig. 1A). We leverage eigenmode representations to model age- and sex-adjusted cortical thickness norms while accounting for site effects. Conventional mass-univariate normative models typically estimate the normative range of a predefined phenotype, such as the average cortical thickness of a fixed region of interest. Here, we term these conventional models "*direct* normative models," with the spatial region of interest referred to as a "*spatial query*." (Fig. 1B; see Section 4.8.1 Spatial Normative Query). The spatial query thus encompasses standard brain-wide or regional averages, while also supporting flexible, weighted queries to characterize fine-grained variations, such as thickness norms in the vicinity of individual vertices.

SNM extends direct normative modeling by embedding cortical measurements into a low-dimensional eigenmode basis. Using connectivity eigenmodes computed via singular value decomposition of a random-walk Laplacian from a high-resolution group-level structural connectivity matrix (Fig. 1C), SNM proceeds in three stages: first, cortical thickness maps are projected onto the eigenmode basis; next, each eigenmode coefficient is modeled with a direct normative model to estimate its normative range while accounting for cross-basis covariance; finally, spectral estimates are combined to generate normative ranges for any spatial query without additional training. In essence, SNM synthesizes pretrained norms of eigenmode coefficients to approximate normative ranges for arbitrary spatial delineations.

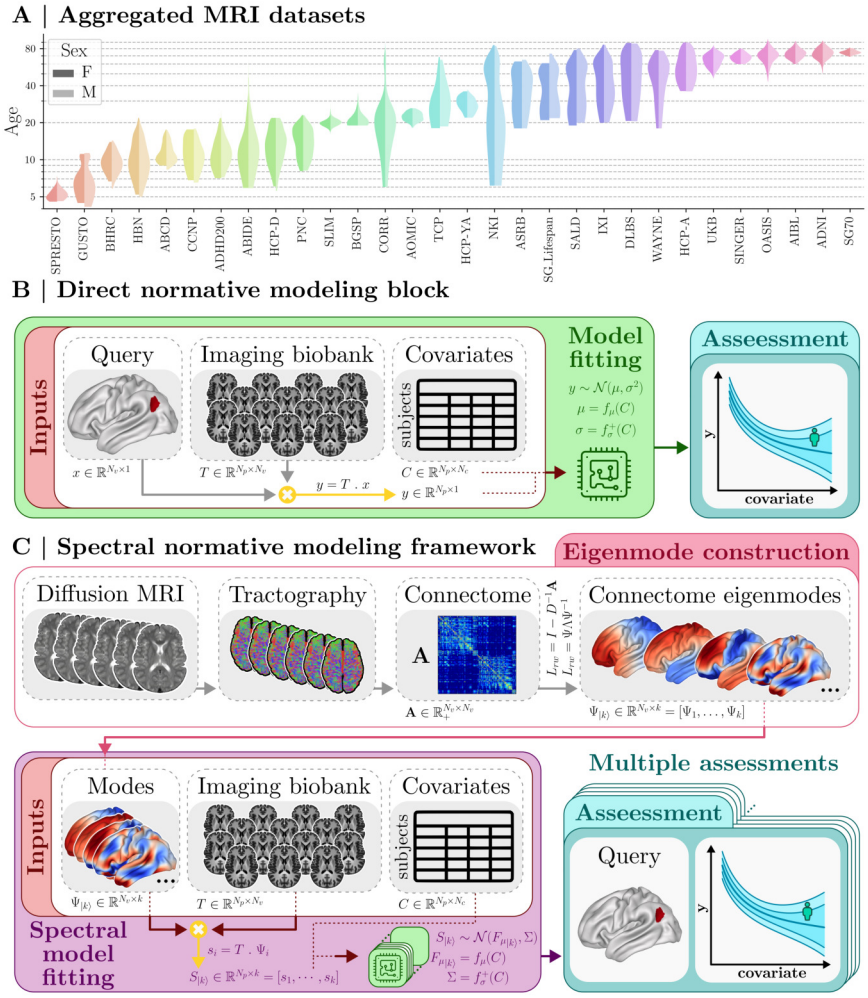


Fig. 1. Spectral Normative Modeling (SNM) to Chart Multi-Scale Lifespan Normative Growth Trajectories. (A) We aggregated quality-controlled, preprocessed MRI data from 78,405 healthy participants across 30 datasets and 189 scanning sites, spanning ages 5-95 years. Violin plots show the age distributions for both sexes in each dataset. Detailed demographic information for all datasets, sites, and scanners is provided in Supplementary Table S.1. **(B)** Architecture of a conventional univariate (direct) normative modeling block, which models the distribution of a phenotype of interest y, obtained by applying a fixed spatial query to imaging data, as a function of covariates C. **(C)** Architecture of the SNM framework, which extends direct normative modeling by removing the need for a predefined spatial query. Using diffusion MRI tractography, we construct a group-level, vertex-resolution structural connectome A capturing the major anatomical pathways linking the cortical sheet. Laplacian eigendecomposition of A yields connectome eigenmodes that serve as a low-dimensional information reconstruction basis. By modeling the multivariate distribution of spectral coefficients $S_{|k\rangle}$ as a function of covariates C, a pretrained SNM (with k modes) can estimate normative ranges for any user-defined spatial query, enabling seamless generation of charts at arbitrary spatial resolutions, including vertex-level maps and multiple cortical parcellations.

2.2 SNM Accurately Charts Population Trajectories Across Spatial Scales

Unlike mass-univariate normative approaches, which are trained on the exact phenotype associated with each spatial query, SNM does not rely on query-specific training. Instead, any query is represented through the multivariate distribution of its spectral coefficients. Each coefficient corresponds to a brain eigenmode, with early modes encoding smooth, global structure and later modes progressively capturing finer, more localized spatial detail (Fig. 2A). As a result, SNM accuracy is governed by two factors: the spatial complexity of the queried phenotype and the dimensionality of the spectral manifold, i.e., the number of eigenmodes retained. To examine their impact, we systematically evaluated performance across multiple spatial queries and spectral dimensionalities by constructing three families of queries (brain-wide, regional, and high-resolution; see Section 4.9 Model Evaluation) and training SNMs with varying numbers of eigenmodes ($k = \{10, 10^2, 10^3, 10^4\}$).

Having established that eigenmodes accurately reconstruct cortical thickness and spatial query information (Extended Data Fig. E.1), we benchmarked SNM against direct normative models trained on the corresponding query-specific phenotypes, with the goal of assessing how closely the spectral approximation recapitulates the direct-model benchmark. Both models were trained on the same cohort of 62,724 healthy participants and evaluated on an independent held-out sample of 15,681 participants, using mean absolute error (MAE) and explained variance (EV) to assess central tendency and mean standardized log-loss (MSLL) to evaluate distributional accuracy[2]. As shown in Fig. 2, SNM accuracy increases monotonically with spectral dimensionality before plateauing at direct-model performance, with the required number of modes determined by the spatial complexity of the query: brain-wide phenotypes are well captured with as few as 100 modes, whereas regional and high-resolution queries require higher-dimensional representations reflecting finer spatial detail encoded in higher-frequency eigenmodes. Importantly, only 1000 modes, less than 2% of the original spatial dimensionality, are sufficient to match direct-model performance across all spatial scales, with additional modes yielding only marginal gains. Notably, training even the most computationally demanding configuration

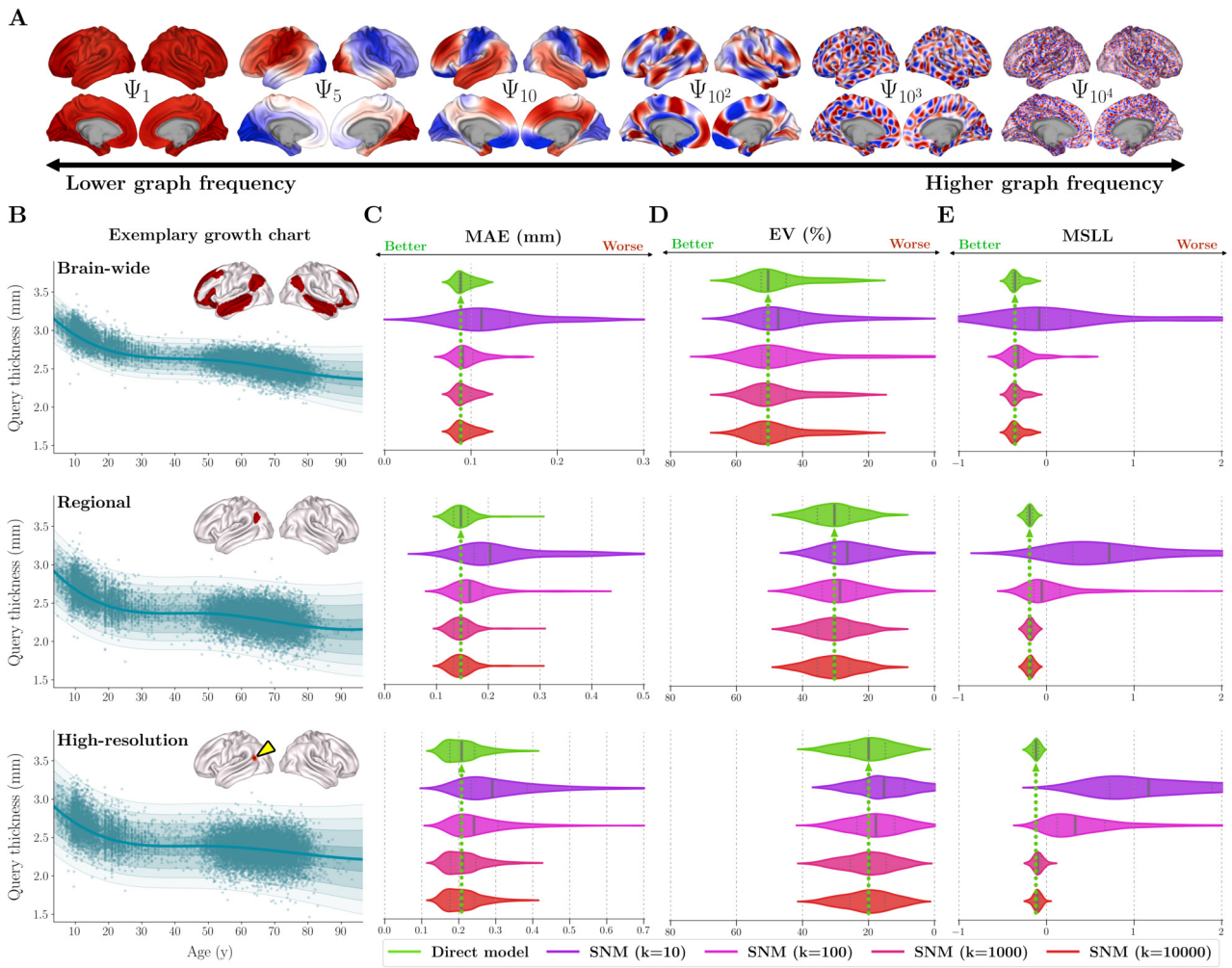


Fig. 2. SNM Normative Performance. This figure benchmarks SNM performance across different numbers of modes ($k = 10, 10^2, 10^3, 10^4$) relative to the direct approach. **(A)** Increasing the number of modes progressively increases spatial specificity: low-order modes capture smooth, global structure, whereas higher-order modes encode finer, more localized patterns. **(B)** Performance is evaluated across three spatial scales of brain-wide, regional, and high-resolution queries, with exemplary normative charts and cortical visualizations illustrating one representative query per scale. Performance metrics include mean absolute error (MAE, **C**) and explained variance (EV, **D**), which quantify accuracy in modeling central tendencies, and mean standardized log-loss (MSLL, **E**), which evaluates the full predictive distribution. For all violin plots, leftward shifts indicate better performance; green violins denote the direct model, and SNM results are shown in shades from purple to red by the number of modes. Violins reflect variability across different queries within each spatial scale. Solid and dashed lines indicate median and first/third quartiles, respectively, with a green arrow marking the direct model's median for reference. Across all evaluations, SNMs with $\geq$1000 modes achieve performance comparable to the direct model.

($k = 10{,}000$) required under 9 hours on a high-performance computing node, highlighting the practical scalability of SNM even at high spectral dimensionalities and sample sizes. Supplementary analyses further characterize regional performance and the effects of query symmetry and granularity (see Supplementary Section S.1.7 Cortical Projections of Performance Metrics and Section S.1.8 Sensitivity Analyses for Query Characteristics).

2.3 SNM Enables High-Resolution Normative Brain Charting

A key advance of SNM is its ability to concurrently generate normative ranges for multiple spatial queries using a single pretrained model (Fig. 3A). By representing cortical phenotypes in a low-dimensional graph–spectral manifold, SNM enables efficient, vectorized inference for many regions of interest, eliminating the need to train separate models per vertex. Using the fitted model with $k = 1000$ modes (SNM-1000), we produce vertex-wise normative charts with an approximately 98% reduction in computation time relative to mass-univariate computations. We leverage this efficiency to derive high-resolution maps of the population mean and variance of cortical thickness, as well as the percent yearly change estimated from the derivative of the normative median.

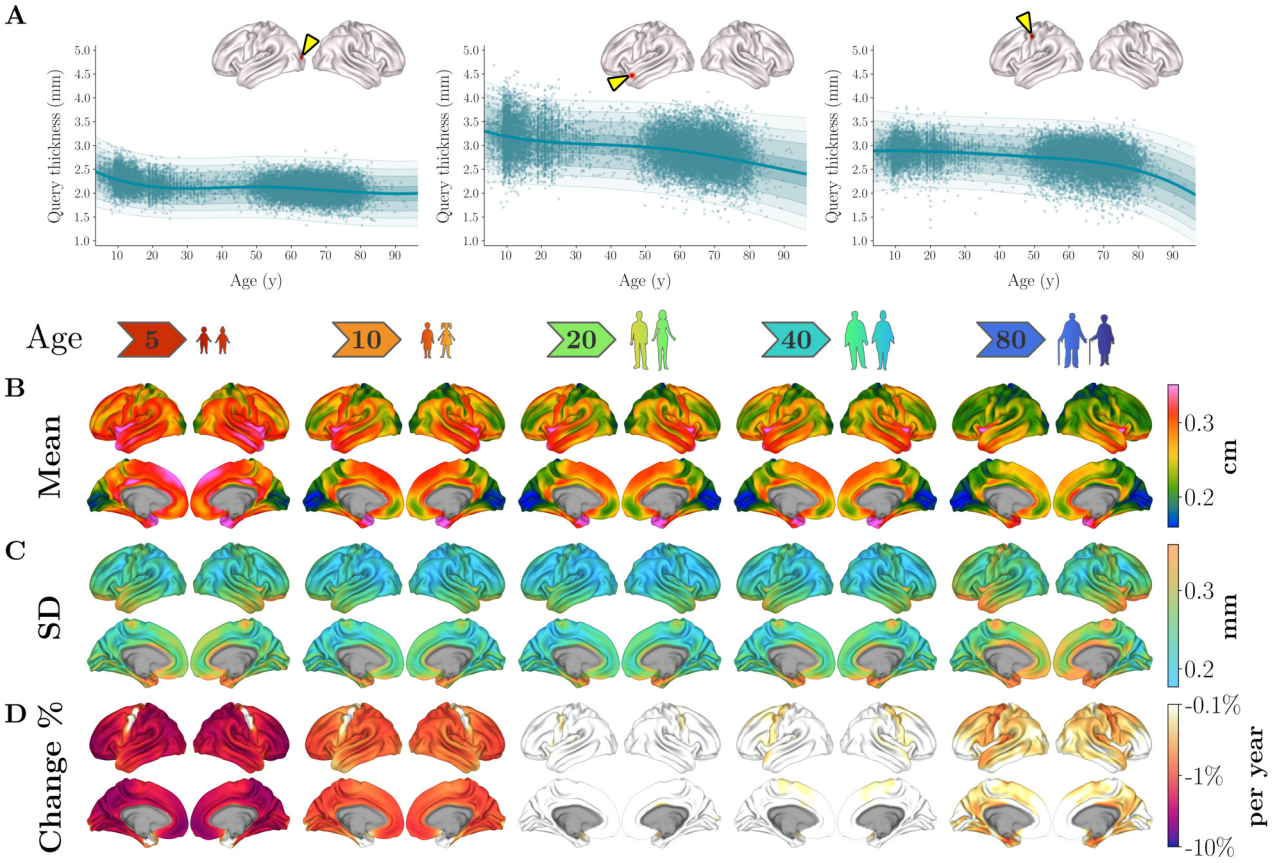


Fig. 3. SNM-Derived High-Resolution Charts of Cortical Thickness across the Lifespan. A single pretrained spectral model enables the extraction of lifespan thickness growth charts for many spatial queries simultaneously. **(A)** Normative trajectories are shown for three example vertices located in occipital, temporal, and precentral cortices. Shaded regions denote the [25–75], [5–95], [1–99], and [0.1–99.9] centile bands, and scatter points represent harmonized individual observations with site effects removed. Extending this procedure to all cortical vertices yields high-resolution maps of cortical thickness development and aging, showing normative patterns of **(B)** mean thickness, **(C)** intersubject variation (distributional spread), **(D)** and percent yearly change at five representative ages (5, 10, 20, 40, and 80 years). The yearly rate of change was computed from the derivative of the median trajectories. All charts were generated using an SNM with 1000 modes.

These charts reveal intricate cortical patterning of thickness across the lifespan (Fig. 3B–D). Mean maps show that regions such as the insula, temporal pole, and middle temporal gyrus remain among the thickest throughout life, whereas occipital cortex consistently exhibits lower thickness. Systematic asymmetries along the central sulcus are also evident, with the precentral gyrus remaining thicker than the postcentral gyrus across all ages, consistent with prior observations [35].

Variance maps reveal patterns of interindividual variability across the lifespan; variability remains relatively stable through adolescence and adulthood but increases in later life, indicative of accumulating divergence in aging trajectories. Higher regional variance is consistently observed near the medial wall, including the parahippocampal and cingulate gyri, likely reflecting a combination of true anatomical heterogeneity in limbic regions [36] and reduced measurement reliability in these areas.

Percent-change maps delineate two periods of accelerated cortical thinning, during development and in late life, separated by relative stability in young adulthood. The spatial patterns of thinning differ across these phases: regions such as the precentral gyrus and temporal pole show minimal thinning during early development but undergo steep declines in late life. These contrasts highlight distinct neurobiological processes underlying cortical remodeling across development and aging[37].

2.4 Gradients of Lifespan Thickness Growth

To characterize organizational principles in normative thickness development and aging, we applied principal component analysis to the derived vertex-wise normative median thickness trajectories, identifying three orthogonal thickness growth gradients (TGG1–3) that together explained over 90% of variance in vertex-wise growth charts (Fig. 4A). Cortical vertices were stratified into strata along each gradient (Extended Data Figs. E.2, E.3, E.4) to generate representative mean and percent-change trajectories, providing continuous profiles of thickness growth variation along gradients (Fig. 4B,C).

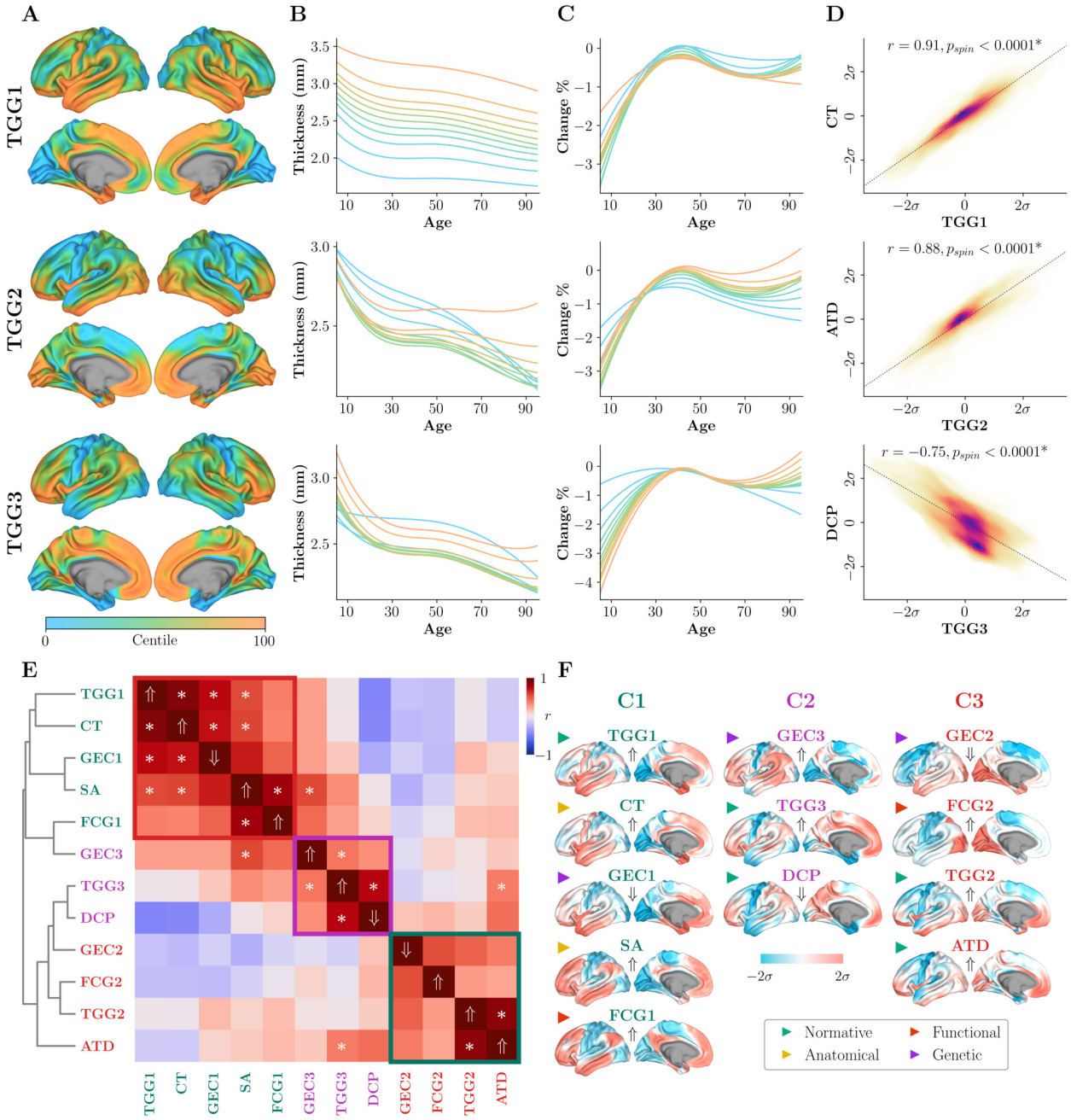


Fig. 4. Lifespan Thickness Growth Gradients. SNM-1000 was used to estimate median cortical thickness at one-year increments from ages 5 to 95. **(A)** Principal component decomposition of these estimates yielded three thickness growth gradients. Cortical projections show the spatial distribution of each gradient, with vertex colors indicating the centile rank of gradient values relative to all vertices in the map. Vertices were stratified into groups along each gradient to define spatial queries, which were used to generate gradient-aligned charts of median thickness **(B)** and percent annual change **(C)**, thereby illustrating the characteristic growth scheme captured by each gradient (see also Extended Data Figs. E.2, E.3, E.4). **(D)** To aid interpretation, gradients are respectively compared against cortical maps of young-adult mean thickness (CT), aging thickness decline (ATD), and developmental cortical pruning (DCP); $p_{spin} < 0.0001$, FDR-adjusted across all 3 tests. **(E)** A data-driven exploratory clustering groups the gradients with distinct anatomical, functional, and genetic hierarchies. The heatmap shows pairwise correlations across cortical maps ($^*p_{spin} < 0.05$, FDR-adjusted across all 66 tests); arrows indicate instances where maps were sign-flipped to maintain positive within-cluster associations (⇑ unflipped, ⇓ flipped). The dendrogram (left) illustrates the hierarchical similarity structure. **(F)** Cortical projections of the sign-flipped maps illustrate cluster-level correspondences. Vertex values of the maps shown in **(D)** and **(E)** were normalized (demeaned and scaled by the reciprocal of the standard deviation) to aid visual comparison. Abbreviations: TGG, Thickness Growth Gradient; CT, Cortical Thickness; ATD, Aging-related Thickness Decline; DCP, Developmental Cortical Pruning; GEC, Gene-Expression Component; FCG, Functional Connectivity Gradient; SA, Sensorimotor–Association axis.

This analysis revealed three distinct, interpretable growth schemes. TGG1 captures a spatially stable hierarchy across the lifespan (Extended Data Fig. E.5), distinguishing regions that remain consistently thick from those that are characteristically thin, effectively reflecting the global pattern of population-wide mean cortical thickness. TGG2 primarily reflects late-life dynamics, with one end encompassing regions that undergo pronounced thinning after mid-adulthood and the other including areas that remain relatively stable in older ages (Extended Data Fig. E.6). TGG3, in contrast, captures early-life developmental effects, separating cortical regions by the rate of thinning during childhood

and adolescence (Extended Data Fig. E.7).

To validate these interpretations, we compared each gradient against cortical maps reflecting the corresponding growth scheme (Fig. 4D). TGG1 strongly corresponded with an established average cortical thickness map[38] (CT; $r = 0.91, p_{spin} < 0.0001$), supporting its role as a global normative thickness axis. TGG2 aligned with a map of aging-related thickness decline (ATD; defined as the average percent change between ages 40 and 95; $r = 0.88, p_{spin} < 0.0001$), while TGG3 corresponded with developmental cortical pruning (DCP; average percent change between ages 5 and 20; $r = -0.76, p_{spin} < 0.0001$). Collectively, these results confirm that the three gradients capture global, late-life, and early-life cortical thickness dynamics. We further observed a negative association between ATD and DCP ($r = -0.43, p_{spin} = 0.017$; Extended Data Fig. E.8), reemphasizing that regions with faster early-life pruning tend to thin less in later life.

Finally, to situate the lifespan thickness growth gradients within established cortical hierarchies, we compared TGG1–3, CT, ATD, and DCP maps with three cortical gene-expression components[39] (GEC1–3), the first two principal gradients of functional connectivity[40] (FCG1–2), and the sensorimotor–association axis[41] (SA). Exploratory data-driven clustering (Fig. 4E) revealed three clusters of spatially similar maps: C1 grouped TGG1 and CT with GEC1, FCG1, and SA; C2 included TGG3 and DCP with GEC3; and C3 comprised TGG2 and ATD with GEC2 and FCG2. While some maps aligned predominantly with a single cluster (e.g., GEC1 with C1), others showed partial alignment across clusters (e.g., GEC1 with C1 and C2), highlighting the interplay of multimodal cortical hierarchies. Many of these associations remained significant under stringent spatial-correspondence testing (two-sided spin-permutations, FDR adjusted across 66 pairwise comparisons), indicating that the normative lifespan thickness growth gradients are closely linked to fundamental anatomical, functional, and transcriptional hierarchies of the human cortex.

2.5 SNM Uncovers Cortical Signatures of Atrophy in Alzheimer's Disease

SNM's pretrained brain charts can be adapted to new imaging sites using only a small healthy sample, enabling translational potential for normative clinical assessment. To illustrate this, we calibrated SNM-1000 on an independent elderly cohort comprising healthy controls (HC, $N = 132$), individuals with mild cognitive impairment (MCI, $N = 202$), and patients with Alzheimer's Disease (AD, $N = 208$; see Supplementary Section S.1.4 MACC Clinical Imaging Data). Fine-tuning harmonization parameters via transfer learning to this dataset allowed identification of vertex-wise deviations in cortical thickness relative to the normative reference (Fig. 5A).

These vertex-resolution z-score maps build on prior regional findings[14,42,43], revealing the characteristic cortical signature of AD atrophy (Fig. 5B), with pronounced normative thinning across temporal, parietal, and frontal regions. In contrast, relative preservation of thickness is observed in parts of the visual cortex, insula, anterior cingulate, and medial orbitofrontal regions. These deviations are linked to cognitive performance, as assessed by Mini-Mental State Examination (MMSE) scores, demonstrating significant associations between cortical thinning and cognitive impairment, strongest in temporal regions, but extending across multiple cortical areas (Fig. 5C). Notably, the spatial pattern of cognitive-impairment–related thinning aligns selectively with the second thickness growth gradient (TGG2; $r = -0.40$, FDR-adjusted spin-test $p = 0.02$), but not with TGG1 or TGG3, indicating that the topography of AD-related atrophy mirrors patterns of accelerated cortical aging[44].

These cortical atrophy maps closely recapitulate well-established anatomical signatures of AD[45–47], validating our normative assessments when adapted to an independent sample. Inspired by prior approaches[14,48], we derive patient-specific abnormality profiles at vertex resolution using extreme value statistics. For each individual, the Extremely Thin Vertex Count (ETVC) quantifies the number of vertices with cortical thickness below a normative threshold (e.g., $z < -1.96$), measuring the extent of severe atrophy. ETVC robustly predicts cognitive impairment across the full sample ($r = -0.53$, $p < 0.0001$), within AD ($r = -0.32$, $p < 0.0001$) and MCI ($r = -0.20$, $p = 0.005$) subgroups (Fig. 5D; Supplementary Section S.1.9 Within-Group Cognitive Associations). Threshold sweeps highlight that vertices with extreme atrophy ($-3 < z < -2$) drive this effect. Supplementary analyses show that ETVC outperforms normative mean cortical thickness in predicting cognitive scores (Supplementary Section S.1.10 Brain-Wide Cognitive Associations), underscoring the utility of high-resolution brain charts for deriving clinically relevant predictive biomarkers.

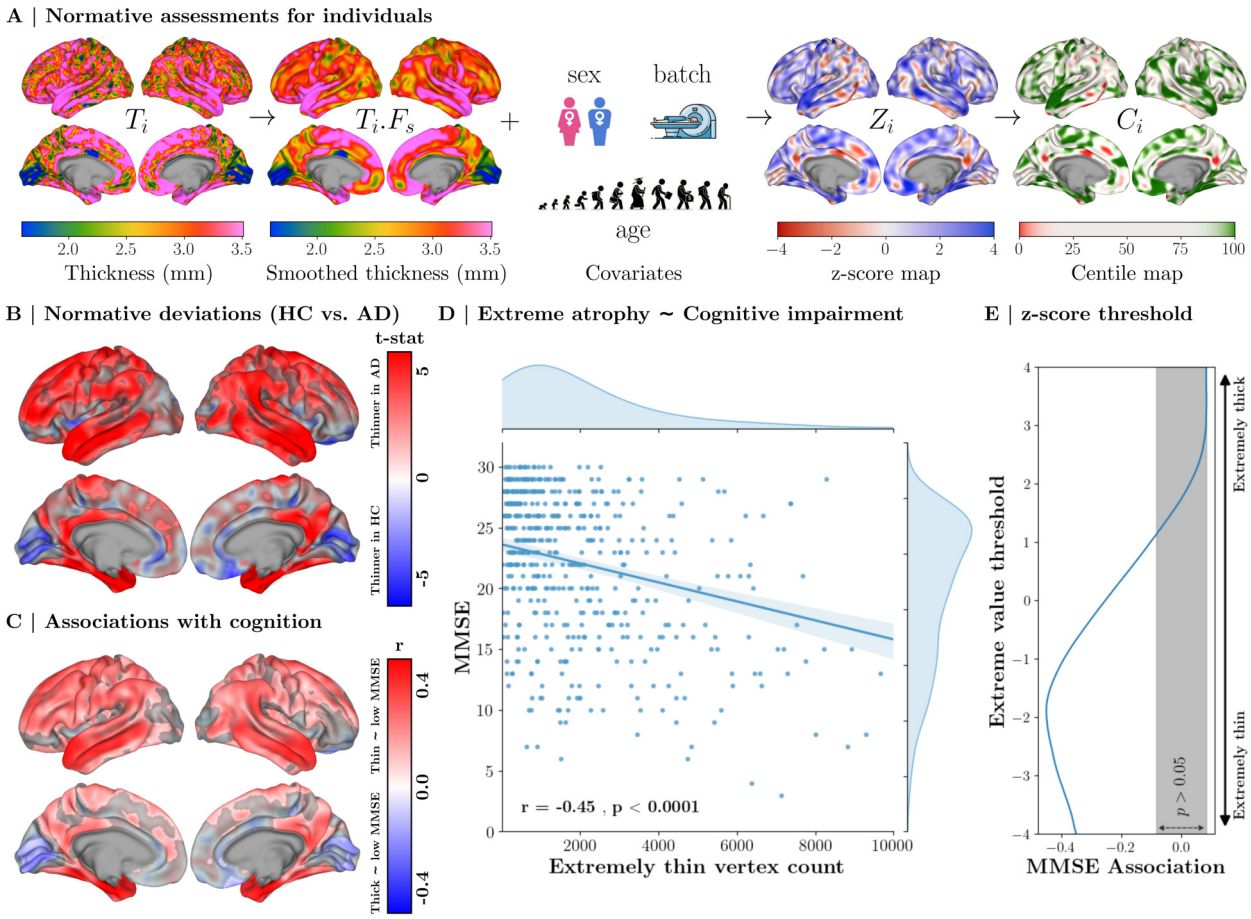


Fig. 5. Cortical Signature of Atrophy in Alzheimer's Disease and Its Cognitive Correlates. Individualized cortical atrophy assessments in AD can be derived from pretrained SNM models, revealing characteristic atrophy patterns and cognitive associations. **(A)** Schematic illustrating SNM-based assessment of individual deviations. Individual cortical thickness maps are smoothed, and vertex-resolution normative deviation maps (z-scores or centiles) are then generated based on the individual's covariates; cortical projections show an exemplary participant. **(B)** Group-level normative differences between HC and AD. **(C)** Vertex-wise associations between normative deviation z-scores and cognitive performance (MMSE). **(D)** The ETVC metric, quantifying extreme atrophy, predicts cognitive performance in the clinical cohort. **(E)** Comparison of z-score thresholds for ETVC reveals that vertices with extreme atrophy provide the highest predictive power for cognitive impairment. Cortical projections highlight significant regions (dimmed for non-significant voxels) at $\alpha = 5\%$ after FDR correction. Abbreviations: HC, Healthy Cohort; AD, Alzheimer's Disease; MMSE, Mini-Mental State Examination; ETVC, Extremely Thin Vertex Count; FDR, False Discovery Rate.

2.6 SNM Highlights Heterogeneity in Individual Cortical Atrophy

While group-level analyses reveal average patterns of AD-related cortical atrophy, they inevitably obscure interindividual variability. Group means can mask highly personalized atrophy patterns, which may hold crucial insights into disease mechanisms and inform individually-tailored clinical interventions. SNM's vertex-resolution assessments capture these individual deviations, enabling detailed exploration of heterogeneity in cortical atrophy among AD patients. As illustrated in Fig. 6A, even individuals with similar demographics and diagnoses can exhibit markedly distinct thinning patterns, with minimal spatial overlap ($r = 0.05$, $p_{spin} = 0.24$). Importantly, these deviation maps can be projected onto each individual's native brain space, in either surface or volumetric format, yielding clinically interpretable atrophy maps that can potentially support personalized evaluation and monitoring.

Differences in atrophy patterns may reflect distinct mechanisms through which pathology associates with normative deviations. To illustrate this, we simulated four hypothetical scenarios (Fig. 6B), each presenting differences across 60 individual normative deviation maps. Pairwise differences were quantified by Euclidean distance to measure interindividual variability. In the first scenario, deviations occur in a random, non-systematic manner. In the second scenario, deviations are tightly linked to diagnostic categories. In the third scenario, deviations are linked to disease severity, with individuals of similar severity exhibiting similar atrophy patterns. In the final scenario, increasing severity produces progressively greater heterogeneity in deviations.

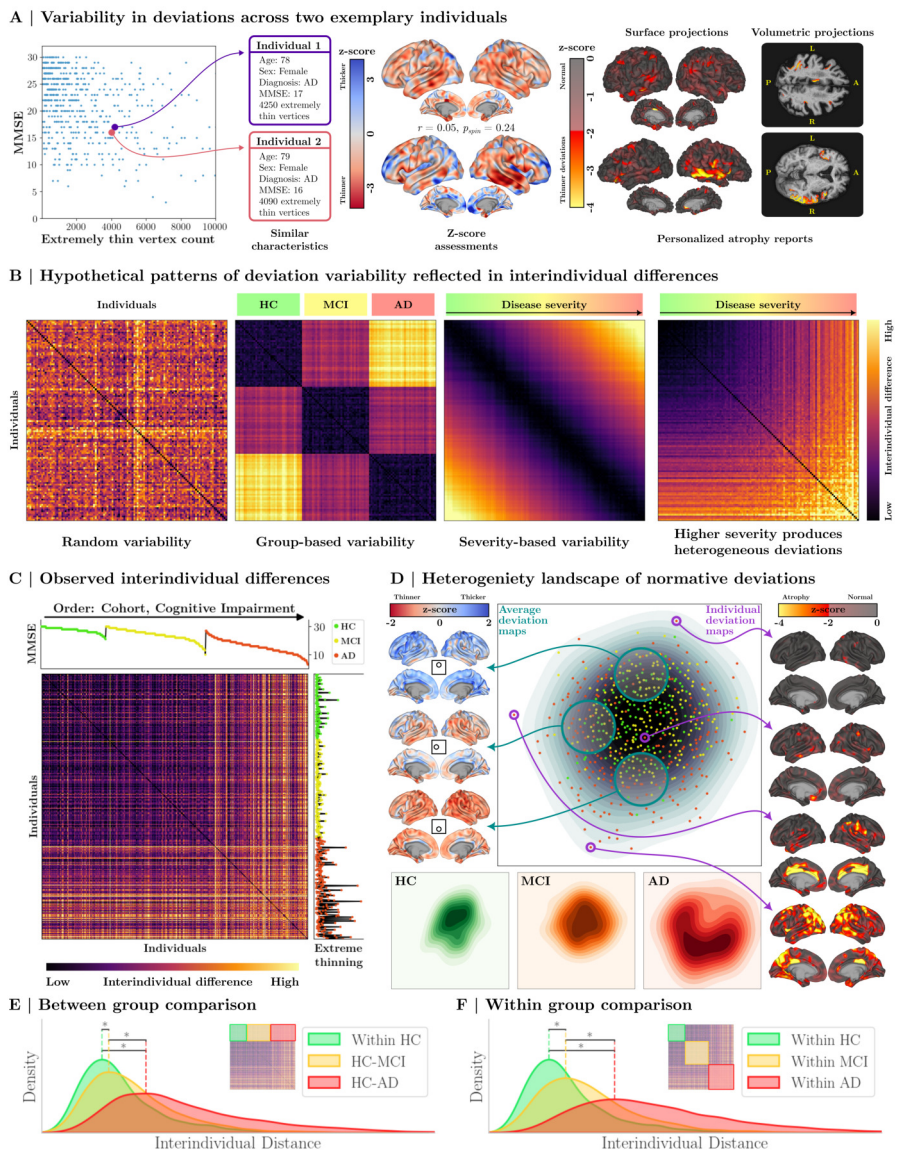


Fig. 6. Heterogeneity Landscape of Atrophy in AD. Individualized deviation maps were used to characterize heterogeneity in AD-related cortical atrophy. **(A)** Two exemplary individuals with similar age, sex, cognitive scores, and diagnosis who exhibit markedly different cortical atrophy patterns. **(B)** Interindividual differences are quantified via Euclidean distances between maps; schematic examples illustrate four hypothetical deviation mechanisms: random deviations unrelated to pathology, diagnosis-aligned deviations, uniformly progressive deviations along disease stages, and severity-induced heterogeneous divergence from norms. **(C)** The empirical interindividual distance matrix, sorted by diagnosis and, within each diagnosis, by decreasing cognitive performance (MMSE), revealing heightened heterogeneity in deviations of AD-diagnosed individuals. **(D)** Deviation maps are embedded into a 2D landscape preserving interindividual differences, with HC (green), MCI (yellow), and AD (red) cohorts occupying partially overlapping regions. Density plots (bottom) highlight regions predominantly occupied by each cohort. Average deviation maps within selected subregions of the landscape (left) and exemplary individual AD maps (right) are shown. **(E–F)** Distributions of empirical interindividual distances across cohorts. **(E)** Between-group comparisons relative to HC show progressively increasing deviations in MCI and AD cohorts. **(F)** Within-group heterogeneity significantly increases from HC to MCI and AD. (∗: Welch's t-tests, FDR-corrected, $p < 0.0001$)

We next compute the empirical interindividual difference matrix (Fig. 6C) and sort it by diagnostic group (HC, MCI, and AD) and, within each group by cognitive performance (as a marker of symptom severity). Expanding earlier regional studies[14,42,43] to vertex resolution, we find that healthy controls and MCI patients exhibit relatively smaller differences in deviation patterns. However, individuals with AD show significantly greater variability, both when compared to healthy and MCI cohorts, as well as amongst themselves. These findings suggest that AD is associated with increasingly divergent cortical atrophy patterns across individuals, highlighting the heterogeneous nature of individual atrophy in AD.

To further elucidate the heterogeneity in deviation patterns, we embed the high-dimensional normative assessments into a lower-dimensional latent space that preserves interindividual differences (Fig. 6D). Healthy individuals form a dense central cluster, characterized by more homogeneous deviation maps. In contrast, the MCI group shows a more dispersed distribution, reflecting increased variability. The AD cohort exhibits the highest degree of deviation from the healthy cluster, with substantial diversity in the direction of deviations across individuals. This underscores various ways in which AD-diagnosed individuals deviate from neurotypical norms. Averaging deviation patterns within local neighborhoods of this landscape (cortical projections in Fig. 6D, left) reveals distinct subgroups characterized by negligible (top), localized (middle), or widespread (bottom) atrophy (see Supplementary Section S.1.12 Heterogeneity Landscape Subgroups). Exemplary cortical projections in Fig. 6D (right) reiterate that AD-diagnosed individuals vary not only in the severity of cortical atrophy but also in the spatial organizations of these deviations. Supplementary analyses confirm that this heterogeneity landscape is robust across dimensionality reduction methods and parameter

settings (see Supplementary Section S.1.13 Robustness of the Heterogeneity Landscape).

We next corroborated these visual patterns with quantitative analyses of the empirical difference matrix. Relative to healthy individuals, both MCI and AD groups exhibited significantly greater deviations, with larger effects in AD (two-sample t-tests: HC–HC vs. HC–MCI, $t = 22.2$, $p < 0.0001$, $d = 0.26$; HC–HC vs. HC–AD, $t = 100.9$, $p < 0.0001$, $d = 1.48$; HC–MCI vs. HC–AD, $t = 99.7$, $p < 0.0001$, $d = 0.85$; Fig. 6E). Within-group comparisons likewise showed increasing heterogeneity from HC to MCI to AD (HC vs. MCI, $t = 35.1$, $p < 0.0001$, $d = 0.43$; HC vs. AD, $t = 125.8$, $p < 0.0001$, $d = 1.27$; MCI vs. AD, $t = 107.8$, $p < 0.0001$, $d = 1.04$; Fig. 6F). Finally, Hartigan's dip statistic[49,50] applied to pairwise differences revealed strongly unimodal distributions in the full sample (Dip $= 0.0003$, $p \approx 1$) and within AD (Dip $= 0.0013$, $p = 0.996$), indicating that heterogeneity varies continuously rather than forming discrete diagnostic clusters or AD subtypes. Together, these findings highlight how group- or subtype-level analyses can obscure meaningful individual variability in cortical atrophy, motivating modeling approaches such as vertex-resolution normative assessment that explicitly capture the personalized aspects of deviation patterns.

3 Discussion

In this manuscript, we introduce Spectral Normative Modeling (SNM), which builds on prior normative approaches by leveraging eigenmode representations as a compact basis for normative inference. Operating in the spectral domain enables efficient generation of normative brain charts across multiple spatial scales, including vertex resolution. This, in turn, allows growth trajectories to be efficiently examined at high resolution, revealing principal organizational hierarchies that capture distinct lifelong, maturational, and aging processes. This further supports individualized, high-resolution assessments of cortical anomalies, providing a potential pathway toward translating normative modeling into precision clinical applications. To facilitate broader adoption, we release the pretrained models alongside a curated software package to facilitate application to new datasets and alternative clinical cohorts.

3.1 Lifespan Gradients of Cortical Growth

A key advance enabled by SNM is its ability to generate normative estimates for novel, *a posteriori*-defined regional delineations without retraining the model. Combined with the computational efficiency afforded by spectral representations, which substantially reduce complexity when concurrently generating tens to hundreds of thousands of normative ranges across queries, SNM facilitates normative mapping at higher resolutions. This capability enabled the characterization of principal gradients of cortical growth across the lifespan. Moreover, by charting normative ranges along gradient strata, SNM provides an interpretable framework for probing neuroanatomical growth schemes along cortical hierarchies.

Our results provide a new perspective on the brain's macroscopic anatomical organization across the lifespan. We show that normative cortical thickness is structured along three distinguishable growth gradients. The first gradient (TGG1) mirrors previously reported maps of mean cortical thickness derived from restricted age ranges[38], indicating that global thickness patterns are largely stable across life. In contrast, the two higher-order gradients (TGG2 and TGG3), which have not been previously characterized, capture distinct aging- and development-specific processes. TGG2 primarily differentiates cortical regions by late-life thinning rates, whereas TGG3 reflects regional differences in cortical pruning during childhood and adolescence.

These growth gradients correspond to multiple organizational hierarchies spanning macroscale function to microscale neurobiology. TGG1 aligns with the SA axis[41], the principal functional connectivity gradient[40], and the principal axis of transcriptional variation[39,51], indicating that higher-order association areas remain consistently thicker than unimodal regions across the lifespan, a feature closely linked to evolutionarily and genetically guided cortical development programs[41,52–54]. TGG2 aligns with the second functional connectivity gradient, revealing that visual cortices are relatively resilient to late-life atrophy while precentral motor areas show greater thinning; TGG2's correspondence with the second transcriptional component, enriched in metabolic processes[39], points to a role of metabolic demand in shaping late-life cortical vulnerability[55,56]. TGG3 co-localizes with the transcriptional component most associated with synaptic plasticity and learning[39], reinforcing its link to

developmental pruning during childhood and adolescence[57]. Notably, the observed anticorrelation between developmental and aging thinning processes (DCP, ATD) demonstrates that regions most stable during development are among those most susceptible to atrophy later in life, providing a nuanced extension of the retrogenesis hypothesis[37,44,55] to normative brain charts.

3.2 Accessibility Benefits

The latent spectral formulation of SNM enables a single, pretrained model to generate normative ranges for a wide variety of queries, including high-resolution queries on different cortical surface templates, regional queries on different cortical surface parcellations[38,58–66], or custom probabilistic regional delineations such as strata along gradients, all without requiring access to the original imaging data used for training. This is particularly valuable because the large-scale datasets used to train population-wide normative models are costly and time-consuming to acquire[67–70], restricting access for researchers interested in novel queries. SNM thus overcomes these key barriers to adoption. To support wider use, we provide the pretrained SNM-1000 model alongside a software package for application to new datasets and phenotypes. For clinical scans acquired on previously unseen scanners, local calibration using a small set of healthy reference scans, commonly available in most clinics (for example, scans acquired to rule out pathology), is sufficient for reliable inference without compromising data privacy.

3.3 Clinical Applications in AD

Neurodegenerative mechanisms, such as amyloid-β plaque deposition and tau-related neurofibrillary tangles, are known to be associated with cortical atrophy in AD[71–74]. As signs of cortical atrophy can appear up to a decade before clinical symptoms[75,76], high-resolution normative assessments can facilitate early detection and characterization of neurodegeneration. In this study, the pretrained SNM model was used to examine the spatial distribution of cortical atrophy in a clinical AD cohort, demonstrating how insights derived from a large healthy reference sample can be extended to smaller clinical datasets[77,78], and providing imaging biomarkers with potential applications in AD research[79].

Our analysis recapitulated hallmark group-level AD cortical atrophy patterns, with strong effects centered on the temporal lobe (temporal pole, superior, middle, and inferior temporal gyri)[45–47,74,80–82], and extending to distributed neocortical regions including the supramarginal and angular gyri[45–47,81], superior parietal lobule, frontal gyri, posterior cingulate cortex[47,82–84], and the precuneus[45–47,74,84]. We also observed less commonly reported anomalies, such as higher cortical thickness in selected visual and orbitofrontal regions[85], potentially reflecting compensatory neurobiological mechanisms[86]. Importantly, AD-related atrophy preferentially aligned with the thickness growth gradient associated with aging-related cortical reorganization (TGG2), consistent with the notion of premature aging: regions most susceptible to thinning during healthy aging exhibit the most pronounced atrophy in AD[44,87,88].

By assessing the occurrence of extreme deviations[48] across thresholds, ETVC analyses demonstrate that the *focal severity* of cortical thinning, rather than its *spatial extent*, is more closely associated with cognitive impairment in AD. Regions exhibiting extreme localized deviations from normative thickness contribute disproportionately to cognitive decline, whereas more widespread but moderate thinning is less informative. Consequently, comparable overall atrophy burdens can have distinct cognitive consequences depending on whether degeneration manifests as localized, severe loss or diffuse, modest reductions. These findings underscore how high-resolution normative frameworks such as SNM are capture cognitively relevant neurodegenerative patterns that are obscured by global atrophy metrics, enabling the detection of early biomarkers before degeneration becomes widespread.

Our evaluations further emphasize the pronounced heterogeneity in individual cortical atrophy patterns in AD, revealing diverse, pathology-driven deviation profiles that are often obscured by group-level analyses[89]. Such heterogeneity may reflect the influence of comorbidities and interacting disease processes[81,90–93], and has traditionally been addressed using clustering approaches that partition patients into putative atrophy subtypes[47,82,94–98]. However, building on recent regional studies of AD heterogeneity[43,99,100], our findings indicate that AD-related cortical deviations are not inherently clusterable. Formal tests of clusterability support the interpretation that these deviation patterns

are more accurately described as a continuous, multidimensional spectrum rather than discrete subtypes[50]. In this setting, imposing subtype boundaries on continuous variation risks obscuring clinically meaningful individual differences and can lead to unstable or imprecise characterizations[101].

Our landscape did not exhibit clear boundaries that would justify subtype separation. Instead, individuals occupied different localities along a continuum: some AD-diagnosed participants showed minimal or negligible atrophy, echoing reports of "minimal-atrophy" subtypes[47,73,82,97], whereas some HC/MCI individuals exhibited marked atrophy without severe cognitive impairment, potentially reflecting resilience mechanisms[102,103]. These observations underscore that modeling deviation patterns along continuous latent dimensions[104] provides a more faithful characterization of neurodegenerative variability than subtype-based categorization.

While these findings highlight the potential clinical utility of SNM for early detection and characterization of atrophy patterns, they remain preliminary, and dedicated application-focused studies will be needed to establish its diagnostic and prognostic value. SNM has potential to support future investigations of differential neurodegenerative patterns across dementia syndromes, including Alzheimer's disease, vascular dementia, Lewy body dementia, and frontotemporal dementia[105]. More broadly, this framework is applicable to other neurological conditions exhibiting localized abnormalities, such as stroke or traumatic brain injury, where deviations could be examined in relation to functional outcomes and recovery trajectories. Finally, SNM's sensitivity to subtle, localized atrophy may facilitate the detection of early biomarkers, such as microinfarcts, enabling identification of at-risk individuals and opening opportunities for preventive intervention before widespread pathology emerges[106].

3.4 Contextualizing SNM in Relation to Deep Learning Approaches

Given the recent emergence of deep learning (DL)–based normative modeling frameworks, often based on variational autoencoders[107–111], it is important to situate SNM within this broader methodological landscape. At a conceptual level, both DL-based models and SNM encode imaging data into a latent space and perform multivariate normative estimation. However, the underlying architectural trade-offs differ substantially. Existing DL-based normative models typically employ deep, nonlinear encoders operating on fixed, low-dimensional regional phenotypes [107–111]. In contrast, SNM operates directly in vertex space with as a wide, single-layer linear encoder defined by fixed eigenmodes rather than trainable weights. This architectural choice enables closed-form normative inference for arbitrary spatial queries, but comes at the cost of reduced capacity to model complex nonlinear relationships. These complementary strengths motivate future development of deep spectral models that integrate eigenmode-based representations with nonlinear learning to achieve both spatial flexibility and increased modeling capacity.

3.5 Future Directions and Limitations

The present study introduces SNM as a flexible framework with scope for continued expansion. Beyond our provided thickness charts, future efforts could extend SNM to additional phenotypes and establish a comprehensive library of spectral normative brain charts. Voxel-wise eigenmodes[112] could extend SNM to whole-brain volumetric phenotypes. Similarly, spectral norms from other modalities, such as volumetric morphometry[13], white matter microstructure[48,113,114], functional organization[115,116], or PET-derived metabolic indices[117], could provide complementary insights. Together, these extensions would contribute to a deeper, multifaceted understanding of neurotypical brain growth.

Several limitations warrant consideration. First, SNM currently assumes Gaussianity, which may limit applicability to phenotypes with skewness or heavy tails; although beyond the scope of this manuscript, likelihood-warping extensions[16,114] could address this limitation. Second, SNM presumes spatial queries lie within a low-pass spectral regime captured by a finite set of eigenmodes. Applications should verify this assumption via reconstruction tests of spatial queries (similar to evaluations in Extended Data Fig. E.1). Finally, although we used random-walk Laplacian eigenmodes, SNM is agnostic to the choice of basis. The software supports alternative bases (e.g., derived from geometry[118], diffusion[25,27], and functional[40,119]) and shift operators (combinatorial, normalized, or fractional Laplacians[120,121]), allowing adaptation across phenotypes, modalities, or spatial priors.

4 Methods

In the ensuing sections, we first describe the brain imaging datasets utilized in this work. We provide a formal description of the high-resolution cortical thickness phenotypes used as input to the normative modeling framework. Next, we introduce the mathematical foundations of SNM, covering core concepts from graph signal processing and normative modeling. Finally, we describe experiments undertaken to evaluate SNM's performance and suitability.

4.1 Brain Imaging Data

This study leverages a large multi-cohort biobank comprising N_p=78,405 healthy participants (after quality control and exclusion criteria) aggregated across 30 publicly available neuroimaging datasets[122–156]. Several datasets include multiple acquisition sites and/or scanners, yielding 189 distinct site/scanner combinations, which were modeled as a categorical site covariate in all normative estimations. Participants spanned ages 4–96 with 50.6% female representation. In addition to the healthy datasets used to train the normative model, we evaluated the model's translational capabilities to real-world data from memory clinics (Memory, Ageing & Cognition Centre at the National University of Singapore: MACC)[157,158]. This data comprised 542 samples (ages 50–91, 61.4% female), including 132 cognitively healthy individuals (24.3%), 202 individuals with mild cognitive impairment (MCI; 37.3%), and 208 individuals diagnosed with Alzheimer's disease (38.4%) based on DSM-IV criteria[158]. While specific datasets differ in acquisition characteristics, all T1-weighted (T1w) structural MRI data underwent cortical thickness extraction using FreeSurfer[159], with additional details regarding acquisition and preprocessing procedures for each dataset documented in the Supplementary Information (see Section S.1.1 Healthy Lifespan Imaging Data).

4.2 Train Test Split

Healthy individuals were stratified into train and test splits to enable validation of the fitted models on unseen data. Namely, 80% (N_{train}=62,724 individuals) of the healthy data were used to train the model, and the remaining 20% (N_{test}=15,681) were used for out-of-sample validation ($N_p = N_{\text{train}} + N_{\text{test}}$). A single split fold was used, with a randomized approach that controlled for covariate distributions of age, sex, and dataset by stratification; this ensured that the covariates had similar distributions across splits[160]. Supplementary Fig. S.2 provides a visual summary of covariate distributions before and after splitting.

4.3 Cortical Thickness

For all cohorts, vertex-wise cortical thickness estimates were derived from preprocessed data. To ensure consistent surface geometry across subjects, an adaptive area–preserving barycentric transformation[161] was used to resample thickness measurements from native FreeSurfer surfaces[162] onto the fs-LR 32k template. This template surface uses 32,492 vertices to model each hemisphere, resulting in an average inter-vertex distance of approximately 2 mm; it includes a total of 59,412 vertices after the exclusion of the medial wall. Vertices are aligned between the left and right hemispheres, enabling inter-hemispheric comparison)[142].

In this space, cortical thickness is represented as a high-resolution vector $T_i \in \mathbb{R}^{1\times N_v}$ ($N_v = 59{,}412$), providing a thickness estimate for each vertex of the ith individual's cortical surface. The collection of these high-resolution thickness features across all participants forms the complete high-resolution thickness data sample $T \in \mathbb{R}^{N_p\times N_v}$ which is used to estimate and evaluate normative ranges of thickness variation across the human lifespan. These data are divided into a training sample $T_{\text{train}} \in \mathbb{R}^{N_{\text{train}}\times N_v}$ used to construct the normative model and an independent test sample $T_{\text{test}} \in \mathbb{R}^{N_{\text{test}}\times N_v}$ used to evaluate the normative model's goodness of fit.

4.4 Normative Covariates

In normative modeling, the aim is to provide a statistical model that describes the distribution of a variable of interest y with respect to a set of covariates C (fixed effects), while potentially accounting for

batch effect Z (random effects)[163]. For simplicity, we consider the covariates of age and sex, although other covariates can be straightforwardly incorporated. Moreover, batch effects were included in the model to facilitate harmonization among the three imaging cohorts (development, young adult, and aging).

4.5 Clinical Fine-Tuning

The normative model, pre-trained on N_{train} healthy individuals, was subsequently fine-tuned for the clinical sample of $N_{MACC} = 542$ individuals using transfer learning. During fine-tuning, the model parameters related to age and sex were kept frozen to maintain consistency with the initial healthy cohort model. However, site parameters for dataset harmonization were fine-tuned on a subsample of 66 healthy individuals (50% of the HC data) from the MACC clinical cohort. The fine-tuned model was then applied to the rest of the clinical sample to generate individualized, high-resolution normative deviation maps. These maps were utilized in normative evaluations of the clinical cohort.

4.6 High-Resolution Connectomes

Our proposed approach uses brain connectivity eigenmodes computed from a reference brain connectivity structure[23]. Connectivity eigenmodes were selected for their ability to span both cortical hemispheres, enabling a cohesive representation of interhemispheric relationships. In contrast, geometric modes analyze hemispheres separately, overlooking interhemispheric interactions. Additionally, as many brain disorders with structural deviations from normative patterns are hypothesized to express spatial signatures aligned with anatomical connectivity, connectome eigenmodes are optimally suited to capture such abnormalities[28–33]. High-resolution structural connectivity data were utilized for this purpose[164].

To this end, we sourced connectomes computed from a tractography pipeline detailed elsewhere[165]. In brief, diffusion-weighted imaging data were used to estimate structural connectivity. Probabilistic tractography was conducted using MRtrix3 to reconstruct whole-brain tractograms (5 million streamlines, with anatomically constrained tractography). Tractography endpoints were used to construct connectomes encoding streamline count at the resolution of the fs-LR 32k template (same space as the thickness information, with $N_v = 59{,}412$ vertices representing network nodes). Individual connectomes were combined to form an estimate of group-level connectivity. Connectome spatial smoothing was performed to account for endpoint inaccuracies and improve connectome reliability[166,167]. This yielded a high-resolution weighted adjacency matrix $\mathbf{A} \in \mathbb{R}_{+}^{N_v \times N_v}$ in which element $A_{i,j}$ denotes the group average strength of structural connectivity between nodes i and j. The eigenmodes resulting from this high-resolution connectome mapping pipeline were previously tested to assess their accuracy in encoding brain signals[23].

4.7 Brain Signal Reconstruction

Graph signal processing enables the study of data encoded on a graph/network structure[168]. As such, considering that the human brain is fundamentally a network structure, graph signal processing can be utilized to design new ways to analyze and study the brain[169]. Here, we utilize dimensionality reduction and signal reconstruction techniques that encode cortical information in a lower-dimensional latent space[170,171]. To this end, we use high-resolution connectomes as a weighted graph $\mathcal{G} : (\mathcal{V}, \mathbf{A})$ where $\mathcal{V} = \{1, 2, \cdots, N_v\}$ is the set of high-resolution cortical nodes (vertices) and $\mathbf{A} \in \mathbb{R}_{+}^{N_v \times N_v}$ is a weighted (non-negative) adjacency matrix. We utilize GSP to model any brain signal $x \in \mathbb{R}^{N_v \times 1}$ that is defined on the set of vertices $\mathcal{V}$. In this context, we model vertex-wise cortical thickness estimates as a brain signal.

The random-walk Laplacian matrix $L_{rw} = I - D^{-1}\mathbf{A}$ is used as the graph shift operator (where $D \in \mathbb{R}^{N_v \times N_v}$ denotes the diagonal strength matrix). This shift operator is diagonalized via singular value decomposition $L_{rw} = \Psi \Lambda \Psi^{-1}$ to compute an orthogonal basis for information reconstruction. Here, $\Psi \in \mathbb{R}^{N_v \times N_v}$ denotes the full set of eigenvectors (also called eigenmodes) corresponding to the left singular vectors of L_{rw}, and $\Lambda \in \mathbb{R}^{N_v \times N_v}$ denotes the diagonal matrix of eigenvalues (singular values) associated with the eigenmodes such that $\Lambda_{i,i} = \lambda_i$ is the ith smallest eigenvalue and $\Psi_i \in \mathbb{R}^{N_v \times 1}$ (ith column of Ψ) refers to its associated eigenmode. The graph Fourier transform (GFT) encodes x

into the graph spectral domain $\tilde{x} = \Psi^{-1}x$, where $\tilde{x} \in \mathbb{R}^{N_v \times 1}$ denotes the encoded signal, and $\tilde{x}_i$ is the ith element of $\tilde{x}$ quantifying the signal loading on the ith eigenmode ($\tilde{x}_i = \Psi_i^{-1}x$). Conceptually, this encoding transforms the brain signal x defined in the spatial domain to a latent embedding $\tilde{x}$ in graph spectral domain. The encoded signal in the graph spectral domain can be mapped back to the spatial domain via the inverse GFT: $x = \Psi\tilde{x} = \Psi\Psi^{-1}x = Ix$. Notably, this transformation, utilizing the full set of eigenmodes, yields an exact reconstruction of any brain signal. Our goal is to use only a fraction of Ψ to derive a low-dimensional representation of brain signals. We selected the random-walk Laplacian shift operator as it yields eigenmodes encoding topological frequency such that eigenmodes associated with smaller eigenvalues capture signals that vary more smoothly along network edges[172].

Deriving the full set of eigenmodes (Ψ_1 to Ψ_{N_v}) is computationally demanding; however, graph Fourier filtering enables efficient approximation of brain signals using a filter that only requires a limited set of eigenmodes. Specifically, we can define a diagonal filtering matrix $H \in \mathbb{R}^{N_v \times N_v}$, where $H_{i,i} = h(\lambda_i)$ denotes the filter's frequency response to the eigenmode Ψ_i associated with the frequency λ_i. An ideal low-pass filter $H_{|k\rangle}$ with frequency response h_k that discards the information of eigenmodes with graph frequencies higher than λ_k is defined using the step function:

$$H_{|k\rangle_{i,i}} = h_k(\lambda_i) \tag{1}$$

$$h_k(\lambda) = \begin{cases} 1 & \text{if } \lambda \leq \lambda_k \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Using this ideal low-pass filter, brain signal x is approximated by $x \approx \hat{x}_{|k\rangle} = \Psi H_{|k\rangle}\Psi^{-1}x = \Psi H_{|k\rangle}\tilde{x}$, where $\hat{x}_{|k\rangle}$ is the low-pass filtered approximation of x reconstructed from the first k eigenmodes. It should be noted that all elements beyond the kth row and column of $H_{|k\rangle}$ are zeros. As a result, this approximation does not require knowledge about higher frequency eigenmodes $\Psi_i : i > k$. Specifically, let $\Psi_{|k\rangle} \in \mathbb{R}^{N_v \times k}$ denote the submatrix of Ψ formed by its first k eigenmodes, $\Psi_{|k\rangle} = [\Psi_1, \ldots, \Psi_k]$, i.e., the first k columns of Ψ. We use $\Psi_{|k\rangle}^{-1} \in \mathbb{R}^{k \times N_v}$ to denote the corresponding submatrix of Ψ^{-1} containing the first k rows, such that $\Psi_{|k\rangle}^{-1}\Psi_{|k\rangle} = I_k$. The low-pass approximation of x is then given by:

$$\hat{x}_{|k\rangle} = \Psi_{|k\rangle}\Psi_{|k\rangle}^{-1}x = \Psi_{|k\rangle}\tilde{x}_{|k\rangle} \tag{3}$$

Where $\tilde{x}_{|k\rangle} \in \mathbb{R}^{k \times 1}$ is the latent spectral representation of x via the first k modes. This low-pass filtering effectively reduces an N_v-dimensional brain signal x to k dimensions. We anticipate that the majority of spatial variation in a brain signal such as cortical thickness is captured by a low-pass approximation of adequate bandwidth (i.e., an appropriate k). The choice of k controls the trade-off between computational complexity and spatial specificity; lower k reduces computation but spatially simplifies the signal; an optimal k balances between computational tractability and spatial specificity.

4.8 Normative Modeling Framework

In the context of our paper, conventional normative models encompass all univariate approaches in which the normative ranges of a fixed spatial query (e.g., the thickness of a region of interest) are modeled as a function of demographic covariates. SNM is a general framework that is compatible with alternative normative models, such as GAMLSS[2,173], BLR[17,174], and HBR[15,163], etc.). For clarity and without loss of generality, we focus on an exemplary instantiation using Hierarchical Bayesian Regression (HBR)[15], although the accompanying software package already supports Bayesian Linear Regression (BLR)[17,174] as an alternative model. We begin by formally defining a spatial normative query to specify the region of interest, followed by an overview of HBR (referred to as the direct method, see Fig. 1A). We then describe how SNM extends the direct approach by using eigenmodes, facilitating the estimation of arbitrary, a posteriori-defined spatial normative queries.

4.8.1 Spatial Normative Query

The concept of a spatial normative *query* is introduced in this work as a mechanism for specifying the spatial extent over which a normative range is to be evaluated. In practical terms, we may wish to

determine the normative range of cortical thickness over a defined region of interest, which can be as fine-grained as a single vertex on the cortical mesh or as broad as a parcel, functional network, or even the entire cortex. Formally, a spatial normative query is represented as a weight vector $x \in \mathbb{R}^{N_v \times 1}$ that linearly maps a cortical phenotype (e.g., a subject's thickness map) to a single summary value via $y = t^\top x$, where $t \in \mathbb{R}^{N_v \times 1}$ denotes the subject-specific thickness measurement.

In its simplest form, a spatial query corresponds to an average over a region of interest. For example, the query defining the average thickness over the entire cortex can be written as $x = \mathbf{1}_{N_v} \cdot \frac{1}{N_v}$, where $\mathbf{1}_{N_v} = [1, 1, \cdots, 1]^\top$ denotes a unit vector of length N_v. Likewise, if the region of interest consists of N_r vertices, we assign uniform weights to those vertices and zero elsewhere, defining the query as

$$x_i = \begin{cases} \frac{1}{N_r} & \text{if vertex } v_i \text{ belongs to the region of interest} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Here, x_i denotes the ith element of $x = [x_1, x_2, \cdots, x_{N_v}]^\top$. This formulation provides a flexible interface through which any region, parcel, or network can be queried by adjusting the weights in x.

While spatial normative queries are often expressed as simple averages, this formulation is not restricted to this case. In general, any weight vector x that encodes a meaningful spatial hypothesis can serve as a query, allowing users to interrogate more complex composite spatial patterns of interest. For example, the normative thickness of a functional network can be queried by assigning weights proportional to the probability that each vertex belongs to that network (with normalization such that $\sum_{i=1}^{N_v} x_i = 1$).

Importantly, this flexibility also enables comparative queries, facilitating the assessment of relative normative variations between regions. For instance, cortical asymmetry (lateralization) can be evaluated by assigning positive weights to vertices in the left hemisphere and negative weights to those in the right hemisphere:

$$x_i = \begin{cases} \frac{1}{N_{\text{left}}} & \text{if } v_i \text{ belongs to the left hemisphere} \\ \frac{-1}{N_{\text{right}}} & \text{if } v_i \text{ belongs to the right hemisphere} \end{cases} \tag{5}$$

where N_{left} and N_{right} denote the number of vertices in each hemisphere.

Importantly, the accompanying software package streamlines the application of spatial queries by providing various built-in utilities, e.g., to project FreeSurfer outputs from native surfaces onto the standard template and to construct queries from arbitrary brain parcellations. This allows users to apply the pre-trained model directly, without having to reimplement preprocessing or query-construction procedures.

4.8.2 Direct Method

The direct method refers to the conventional normative modeling approach in which a fixed a-priori-defined spatial query is used (e.g. mean thickness). In this manuscript, we utilize HBR as the normative approach used for the direct method[15,163], allowing modeling of site variability as random effects via partial pooling and reuse of site hyperpriors to support normative inference at unseen sites. Using a fixed spatial query x, the direct method models the expression of the phenotype (e.g. cortical thickness) on the query in the training set, $y \in \mathbb{R}^{N_{\text{train}} \times 1}$; which is computed as the inner product between the spatial query x and the high-resolution thickness data T, $y = T_{\text{train}} \cdot x$. HBR assumes that these observations follow a normal distribution $y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ and aims to model its moments (mean and variance) as a function of covariates (age and sex) while controlling for batch effects:

$$\mu_y = f_\mu(\text{age, sex, batch}) = w_\mu \Phi(\text{age}) + \alpha_\mu(\text{sex}) + \beta_\mu(\text{batch}) + \epsilon_\mu \tag{6}$$

$$\sigma_y = f_\sigma^+(\text{age, sex, batch}) = w_\sigma^+ \Phi(\text{age}) + \alpha_\sigma^+(\text{sex}) + \beta_\sigma^+(\text{batch}) + \epsilon_\sigma \tag{7}$$

In particular, w_μ and w_σ^+ are weights that model the fixed effect of age by a basis expansion Φ. A B-spline basis expansion (cubic spline with three evenly spaced knots; df=5) was used to capture the non-linear effects of age. The fixed effect of sex is modeled by intercepts for mean (α_μ) and deviation (α_σ^+). Batch effects (for dataset harmonization) are modeled as intercepts for the mean

($\beta_\mu \sim \mathcal{N}(\mu_{\beta_\mu}, \sigma_{\beta_\mu})$) and deviation ($\beta_\sigma^+ \sim \mathcal{N}(\mu_{\beta_\sigma^+}, \sigma^2_{\beta_\sigma^+})$) that are randomly drawn from another prior distribution; this results in hierarchically modeling batch as a random effect. The prior assumptions of the Bayesian model can be summarized as follows:

$$w_\mu(i) \sim \mathcal{N}(0, \sigma^2_{w_\mu(i)}) \quad , \quad \alpha_\mu \sim \mathcal{N}(0, \sigma^2_{\alpha_\mu}) \quad , \quad \beta_\mu \sim \mathcal{N}(\mu_{\beta_\mu}, \sigma^2_{\beta_\mu}) \tag{8}$$

$$w_\sigma^+(i) \sim \mathcal{N}^+(0, \sigma^2_{w_\sigma^+(i)}) \quad , \quad \alpha_\sigma^+ \sim \mathcal{N}^+(0, \sigma^2_{\alpha_\sigma^+}) \quad , \quad \beta_\sigma \sim \mathcal{N}^+(\mu_{\beta_\sigma}, \sigma^2_{\beta_\sigma}) \tag{9}$$

Where $\mathcal{N}^+$ denotes a positive half-normal distribution. The random effect parameters are hierarchically sampled from a set of batch hyperparameters:

$$\mu_{\beta_\mu} \sim \mathcal{N}(0, \sigma^2_{\mu_{\beta_\mu}}) \quad , \quad \sigma_{\beta_\mu} \sim \mathcal{N}^+(0, \sigma^2_{\sigma_{\beta_\mu}}) \quad , \quad \mu_{\beta_\sigma} \sim \mathcal{N}^+(0, \sigma^2_{\mu_{\beta_\sigma}}) \quad , \quad \sigma_{\beta_\sigma} \sim \mathcal{N}^+(0, \sigma^2_{\sigma_{\beta_\sigma}}) \tag{10}$$

Notably, the model can quantify the normative deviation of an individual subject y_i using a z-score:

$$z = \frac{y_i - \mu_y}{\sigma_y} \tag{11}$$

This z-score can be used to compute normative centiles using the cumulative distribution function (CDF) of the normal distribution.

4.8.3 Spectral Normative Model (SNM)

Next, we explain the theory behind the proposed SNM framework. In the training phase, k direct normative models are independently fitted to the projection of phenotypes on the first k brain eigenmodes with the lowest graph frequencies. This forms the low-dimensional latent approximation containing information about the low-frequency characteristics of the cortical phenotype. For example, in the case of cortical thickness, for each $i \leq k$, the spectral coefficient $s_i \in \mathbb{R}^{N_{\text{train}} \times 1}$ associated with Ψ_i is first computed:

$$s_i = T_{\text{train}} \cdot \Psi_i \tag{12}$$

Similar to the direct model, normative estimators of the mean (f_{μ_i}) and deviation ($f^+_{\sigma\ i}$) are trained to model normative ranges of all low-frequency spectral coefficients:

$$\forall i \leq k : s_i \sim \mathcal{N}(\mu_{s_i}, \sigma^2_{s_i}) \Rightarrow \begin{cases} \mu_{s_i} = f_{\mu_i}(\text{age, sex, batch}) \\ \sigma_{s_i} = f^+_{\sigma\ i}(\text{age, sex, batch}) \end{cases} \tag{13}$$

The normative model for each spectral coefficient is the same as the direct case (see equations 6, 7, 8, 9, 10). Next, the normative ranges of any arbitrary spatial query is approximated using this set of k pre-trained normative models. For any spatial query x, its observed cortical phenotype $y = T_{\text{train}} \cdot x$ is approximated by a low-pass graph filter (see equation 3):

$$\begin{aligned} y = T_{\text{train}} \cdot x \approx \ & \hat{y}_{|k\rangle} = T_{\text{train}} \cdot \hat{x}_{|k\rangle} \\ = \ & T_{\text{train}} \Psi_{|k\rangle} \Psi^{-1}_{|k\rangle} x = T_{\text{train}} \Psi_{|k\rangle} \tilde{x}_{|k\rangle} \\ = \ & \sum_{i \leq k} T_{\text{train}} \Psi_i \tilde{x}_i = \sum_{i \leq k} s_i \tilde{x}_i = S_{|k\rangle} \cdot \tilde{x}_{|k\rangle} \end{aligned} \tag{14}$$

where $S_{|k\rangle} \in \mathbb{R}^{N_{\text{train}} \times k}$ denotes the matrix containing the first k spectral coefficients. The variable of interest y can hence be approximated by a linear combination of the spectral coefficients $S_{|k\rangle}$. Notably, if the spatial query is frequency-bounded this approximation will be exact (i.e. if the query is a linear combination of low-frequency eigenmodes such that $\tilde{x}_i = 0$ for $i > k$). Equation 14 indicates that for any spatial query x, the associated phenotype y can be modeled as a multivariate normal distribution such that its normative ranges are estimated by the linear combination $y = S_{|k\rangle} \cdot \tilde{x}_{|k\rangle}$. This is formalized by the following equation:

$$\forall x \in \mathbb{R}^{N_v \times 1},\ y = T_{\text{train}}\ .\ x:\ y \sim \mathcal{N}(\mu_y, \sigma_y^2) \Rightarrow \begin{cases} \mu_y \approx g_\mu(F_{\mu|k\rangle}, \tilde{x}_{|k\rangle}) \\ \sigma_y \approx g_\sigma^+(, \tilde{x}_{|k\rangle}) \end{cases} \tag{15}$$

Where g_μ and g_σ^+ are functions that respectively approximate the mean and standard deviation of the multivariate normal distribution based on the trained low-pass spectral moments $(F_{\mu|k\rangle},)$. Explicitly, g_μ defines the mean of the normal distribution y as a linear combination of the means of each spectral coefficient and is based on the graph spectral encoding $\tilde{x}_{|k\rangle}$:

$$g_\mu(F_{\mu|k\rangle}, \tilde{x}_{|k\rangle}) = \sum_{i \leq k} f_{\mu_i} \tilde{x}_i = \begin{bmatrix} f_{\mu_1}, & f_{\mu_2}, & \cdots, & f_{\mu_k} \end{bmatrix}_{1\times k} .\ \tilde{x}_{|k\rangle} = F_{\mu|k\rangle}\ .\ \tilde{x}_{|k\rangle} \tag{16}$$

where $F_{\mu|k\rangle}$ is used to summarize the set of estimated means for the first k spectral coefficients in $S_{|k\rangle}$ into a $1 \times k$ vector. Note that equation 16 arises because the expected value of a linear combination of a set of random variables is equal to the linear combination of their respective expected values:

$$E[a_1 X_1 + a_2 X_2 + ... + a_k X_k] = a_1 E[X_1] + a_2 E[X_2] + ... + a_k E[X_k] \tag{17}$$

Moreover, g_σ^+ defines the standard deviation of the normal distribution y as a function of the covariance matrix describing cross-basis dependencies of spectral coefficients $\in \mathbb{R}_+^{k\times k}$:

$$g_\sigma^+ = \sqrt{\sum_{i,j \leq k} \tilde{x}_i \tilde{x}_{j\,i,j}} = \sqrt{\tilde{x}_{|k\rangle}^\top \tilde{x}_{|k\rangle}} \tag{18}$$

Equation 18 describes how the standard deviation of the variable of interest y is related to the covariance matrix describing dependencies across spectral coefficients s_i. This equation is based on the following notion that relates the variance of a linear combination of a set of normally distributed variables to their covariance structure:

$$\text{Var}\left(\sum_{i=1}^{k} a_i X_i\right) = \left(\sum_{i=1}^{k} a_i^2 \text{Var}(X_i)\right) + \left(\sum_{i=1}^{k} \sum_{j=1(j\neq i)}^{k} a_i a_j \text{Cov}(X_i, X_j)\right) \tag{19}$$

Elements of can also change as a function of covariates and batch effects. We note that for a pair of random variables (X and Y), covariance ($\text{Cov}(X, Y)$) and correlation ($\text{Corr}(X, Y)$) are related by $\text{Cov}(X, Y) = \sigma_X \sigma_Y \text{Corr}(X, Y)$, where σ denotes the standard deviation of each variable. We use this notion to separate the effect of within-mode variance from the effect of cross-mode correlation on . In particular, the following equation decomposes the cross-basis covariance matrix into a correlation matrix $\mathbf{P}$ and an independent set of eigenmode standard deviations:

$$= \text{diag}\left(F_{\sigma\ |k\rangle}^+\right)\ .\ \mathbf{P}\ .\ \text{diag}\left(F_{\sigma\ |k\rangle}^+\right) \tag{20}$$

where $F_{\sigma\ |k\rangle}^+$ summarizes the set of estimated standard deviations of the first k spectral coefficients $(y_1, \cdots, y_k)$ to a $1 \times k$ vector, and the *diag()* operator is used to map a $1 \times k$ vector to a $k \times k$ diagonal matrix:

$$\text{diag}\left(F_{\sigma\ |k\rangle}^+\right) = \begin{bmatrix} f_{\sigma\ 1}^+, & 0, & \cdots, & 0 \\ 0, & f_{\sigma\ 2}^+, & \cdots, & 0 \\ \vdots, & \vdots, & \ddots, & \vdots \\ 0, & 0, & \cdots, & f_{\sigma\ k}^+ \end{bmatrix} \tag{21}$$

Thus far, we show that any spatial query can be represented as a multivariate normal distribution across k spectral components, the normative ranges of which can be approximated from the mean ($F_{\mu|k\rangle}$), standard deviation ($F_{\sigma\ |k\rangle}^+$), and cross-correlation ($\mathbf{P}$) of spectral coefficients. Mean and standard deviation estimates are computed similarly to a direct HBR model. However, approximations may be needed to estimate the cross-correlation structure. This is because the complexity of the correlation structure increases quadratically with the number of eigenmodes included in the low-pass approximation (k). We hence impose a sparsity constraint on the correlation structure using a

threshold on the observed cross-correlation of spectral phenotypes within the training sample. Our evaluations indicate that the cross-correlation matrix is naturally sparse (see Supplementary Fig. S.3). As a result, the majority of elements $\mathbf{P}_{i,j} : \text{Corr}(s_i, s_j) \leq \rho$ are small and will hence be replaced with zero through this sparsification step. The remaining pairs of suprathreshold correlations form a sparse matrix representation as summarized by the following derivation of $\mathbf{P}$:

$$\mathbf{P}_{i,j} = \begin{cases} 0 & \text{if } \text{Corr}(s_i, s_j) \leq \rho \\ \rho_{s_i,s_j} & \text{if } \text{Corr}(s_i, s_j) > \rho \end{cases} \tag{22}$$

The suprathreshold elements ρ_{s_i,s_j} are affected by covariates (age and sex) and batch effects. As such, for every suprathreshold element, the spectral coefficient pair was modeled as a bivariate normal distribution:

$$\begin{bmatrix} s_i \\ s_j \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_{s_i} \\ \mu_{s_j} \end{bmatrix}, \begin{bmatrix} \sigma^2_{s_i} & \rho_{s_i,s_j}\sigma_{s_i}\sigma_{s_j} \\ \rho_{s_i,s_j}\sigma_{s_i}\sigma_{s_j} & \sigma^2_{s_j} \end{bmatrix} \right) \Rightarrow \begin{cases} \text{constants} \begin{cases} \mu_{s_i} = f_{\mu_i};\ \mu_{s_j} = f_{\mu_j} \\ \sigma_{s_i} = f^+_{\sigma\ i};\ \sigma_{s_j} = f^+_{\sigma\ j} \end{cases} \\ \\ \rho_{s_i,s_j} = tanh\left(f_{\rho_{i,j}}(age,\ sex,\ batch) \right) \end{cases} \tag{23}$$

where spectral coefficients' estimates of the mean (μ_{s_i}, μ_{s_j}) and deviation $(\sigma_{s_i}, \sigma_{s_j})$ were inputted as constants (yielded from the solution of Equation 13). The cross-correlation estimate (ρ_{s_i,s_j}) was adjusted by a hyperbolic tangent function to limit the range of possible correlation values to $[-1, 1]$. Effects of covariates and batch were modeled as follows:

$$f_{\rho_{i,j}}(age,\ sex,\ batch) = w_\rho \Phi(age) + \alpha_\rho(sex) + \beta_\rho(batch) + \epsilon_\rho \tag{24}$$

with the following prior assumptions:

$$w_\rho(i) \sim \mathcal{N}(0, \sigma^2_{w_\rho(i)}) \quad , \quad \alpha_\rho \sim \mathcal{N}(0, \sigma^2_{\alpha_\rho}) \quad , \quad \beta_\rho \sim \mathcal{N}(\mu_{\beta_\rho}, \sigma^2_{\beta_\rho}) \tag{25}$$

and the following hierarchical batch priors:

$$\mu_{\beta_\rho} \sim \mathcal{N}(0, \sigma^2_{\mu_{\beta_\rho}}) \quad , \quad \sigma_{\beta_\rho} \sim \mathcal{N}^+(0, \sigma^2_{\sigma_{\beta_\rho}}) \tag{26}$$

In summary, Equations 12–26 formalize how SNM derives normative ranges for a target phenotype y. The method first projects the cortical data for a given spatial query $\tilde{x}_{|k\rangle}$ into the low-dimensional spectral domain, then models the resulting spectral coefficients $S_{|k\rangle}$ as a multivariate normal distribution. Normative means and deviations for the target phenotype are then estimated directly from this distribution, enabling rapid inference for arbitrary spatial queries without requiring direct training on novel spatial queries.

4.9 Model Evaluation

We use three different families of spatial normative queries (x) to evaluate SNM's performance. These families include (I) *brain-wide*, (II) *regional*, and (III) *high-resolution* signals, each of which groups queries according to their spatial scale. Brain-wide queries are regional masks that describe coarse and large-scale characteristics along the cortex. Namely, this includes a total of 25 queries that represent the average of total cortical thickness, as well as the average thickness over predefined brain networks that divide the brain into 7/17 functionally distinct segments (Yeo functional networks)[61]. Regional queries included 200 signals, each describing the average cortical thickness within a region of interest as defined by a brain parcellation. For regional queries, we used a homotopic parcellation of the cortex comprising 200 cortical regions[175] (Yan200 atlas). In the supplementary analysis, we evaluate the sensitivity of our findings to atlas granularity (see Supplementary Figs. S.10, S.11). Finally, a total of 4,000 high-resolution queries were constructed (2,000 per hemisphere), each describing the average cortical thickness in the vicinity of a selected cortical vertex. Vertices were randomly sampled to ensure uniform coverage across all cortical regions (see Supplementary Section S.4). This subsample is used solely for evaluation purposes; SNM itself scales efficiently to full vertex resolution (i.e., hundreds

of thousands of vertices). The limiting factor in our benchmark design is the direct model comparison, which requires fitting a new model for each query, making large-scale comparison computationally prohibitive.

Every high-resolution query was formed by an 8mm FWHM smoothing kernel based on the geodesic distance to the respective vertex[167]. The smoothing kernel serves two key purposes: (1) to mitigate the impact of local intersubject registration misalignments and (2) to ensure vertex-resolution queries are appropriately reconstructed by the low-pass graph filter, thereby reducing ringing effects in reconstructions. In a supplementary analysis, we evaluate the sensitivity of our evaluations to the smoothing kernel choice (see Supplementary Figs. S.10, S.12). The signals evaluated here quantify the average thickness of the cortex with different spatial granularities (ranging from tens of centimeters to a few millimeters). While such average queries are the only type of normative queries that have been previously studied for brain charting, in supplementary analyses we evaluate the applicability of our proposed model to other possible cases, e.g. when a comparative asymmetric norm is being studied (see Supplementary Figs. S.13, S.14, S.15, and S.16).

For all signal families (whether presented in the main text or supplementary information), we evaluated the performance of SNM (using different values of k) based on (i) reconstruction accuracy, normative model's (ii) goodness of fit of the central tendency, and (iii) total goodness of fit of the estimated distribution. These evaluation criteria are detailed in the ensuing sections.

4.9.1 Reconstruction Accuracy

For the spectral approach to provide accurate estimates of normative ranges, both the cortical phenotype under study (i.e. cortical thickness) and the spatial normative query must be adequately reconstructed by the low-dimensional latent space defined by the brain eigenmodes. We hence use the following two metrics to evaluate the validity of the brain eigenmodes for reconstructing brain signals:

Energy Proportion: For any arbitrary brain signal $x \in \mathbb{R}^{N_v \times 1}$ with graph spectral encoding $\tilde{x} = \Psi^\top x$, the total energy of the signal can be quantified by $E(x) = \sum_{i \in \{0 \cdots N_v\}} x_i^2 = \sum_{i \in \{0 \cdots N_v\}} \tilde{x}_i^2 = E(\tilde{x})$. In other words, the signal energy in vertex domain is equal to the signal energy in the spectral domain (Parseval's theorem[176]). As such, the proportional contribution of every single eigenmode Ψ_i to the total signal energy can be quantified by $\tilde{x}_i^2 / E(x)$. We provide a cumulative line plot of observed ranges for this proportional energy for different brain signals. This indicates the number of eigenmodes that capture major energy proportions for different brain signals.

Standardized Mean Squared Error (SMSE): This measure normalizes the mean squared error (MSE) by signal variance to provide an error metric that does not depend on signal variance[177]. Using this metric, a trivial reconstruction that replaces values of all vertices x_i with the signal mean will have an SMSE of one, and a perfect reconstruction will result in an SMSE of zero. By considering $\hat{x}_{|k\rangle}$ to be a reconstruction of the original brain signal x, we compute SMSE for different ks to quantify reconstruction error at different low-pass frequency bandwidths. A high SMSE (close to one) indicates poor reconstruction performance and a low SMSE (close to zero) marks accurate reconstruction.

4.9.2 Model's Central Tendency

Independent of how accurately the eigenmode basis reconstructs various brain signals, we also need to validate that SNM yields accurate normative fits for spatial queries. This is quantified by measures of goodness of fit of the central tendency[2]. These measures can particularly indicate the accuracy of the normative model in explaining trends in the data. To this end, we utilize the mean absolute error (MAE) and explained variance (EV) metrics. An out-of-sample MAE score was evaluated using 20% of the data in the test set. Specifically, for any spatial query x, the thickness of that query in the test set is measured $y \in \mathbb{R}^{N_{\text{test}} \times 1} = T_{\text{test}}$. x and MAE quantifies the extent to which these thickness values deviate from model predictions $\text{MAE} = \mathbb{E}[|y - \mu_y|] = \sum |y_i - \mu_{y_i}| / N_{\text{test}}$. In addition to MAE, we also quantify the out-of-sample explained variance (EV), which measures the proportion of phenotype variance in the test sample that is captured by the model predictions, $\text{EV} = 1 - \text{Var}(y - \mu_y)/\text{Var}(y)$. MAE and EV were computed for each of the three families of spatial queries, comparing

the performance of a direct HBR model against four alternative SNMs, which utilized the first 10, 100, 1,000, and 10,000 eigenmodes, respectively. Since all models provided mean thickness estimates in millimeters, MAE offers a practical measure of the average error in millimeters when inferring the thickness of a given spatial query.

4.9.3 Model's Total Goodness of Fit

Measures of central tendency, such as mean or median, assess only the accuracy of modeling normative trends but do not capture the validity of the estimated distribution's deviations. To evaluate the model's overall goodness of fit—including both mean and variance estimates—we utilize the mean standardized log loss (MSLL)[177]. Mean log loss quantifies the mean negative log probability of the observed data (in the test sample) under the hypothesis of the fitted normative model:

$$\text{log-loss(model)} = -\frac{1}{N_{\text{test}}} \log\left(p(y|\mu_y, \sigma_y)\right) = -\frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \left(\frac{1}{2} \log(2\pi\sigma_{y_i}^2) + \frac{(y_i - \mu_{y_i})^2}{2\sigma_{y_i}^2} \right) \tag{27}$$

MSLL is computed by comparing the log loss of competing models against the log loss of a trivial reference model that is based on the mean and standard deviation of the training sample, MSLL = log-loss(model) − log-loss(reference). By definition, the trivial model achieves an MSLL of zero, and models providing better than trivial normative ranges will yield negative MSLL values. Here, we estimate log loss using a censored log-likelihood measure to reduce the sensitivity to outliers[2,178]. As such, for observations falling outside of the range of 1st to 99th percentile, the value of log loss was replaced by $-\log(0.02)$. This censored MSLL score favors models with accurate distribution estimates within this percentile range and is thus robust against extreme outliers.

4.10 Mapping Cortical Growth Gradients

The cortical growth gradients reported in Section 2.4 Gradients of Lifespan Thickness Growth were derived using a principal component analysis (PCA) framework applied to normative lifespan predictions from SNM. Specifically, the pretrained SNM with $k = 1000$ eigenmodes was used to estimate median cortical thickness at all $N_v = 59{,}412$ vertices across 1-year age intervals spanning 5 to 95 years, while partialling out site and sex effects. This procedure yielded a matrix $G \in \mathbb{R}^{N_v \times 91}$ containing vertex-wise normative thickness trajectories across the lifespan. To ensure comparability across vertices while preserving relative biological variability, G was globally standardized by subtracting its mean and dividing by its standard deviation, $\tilde{G} = (G - \mu_G)/\sigma_G$. This normalization retains relative magnitude differences across cortical regions, allowing vertices with greater lifespan variability to contribute proportionally to the extracted components.

Centered PCA was then applied to $\tilde{G}$ to identify orthogonal axes of maximal variance. The first three principal components, termed thickness growth gradients (TGGs), were retained and mapped back onto the cortical surface to represent dominant spatial modes of lifespan cortical growth. The first component (TGG1) captured a global thickness factor closely aligned with mean cortical morphology and accounted for 97.4% of the total variance. To examine more subtle spatial patterns beyond this global effect, we analyzed the second and third components (TGG2 and TGG3), which together explained 91.5% of the residual variance not captured by TGG1 (66.5% by TGG2 and 25.0% by TGG3). Conceptually, this procedure summarizes shared and divergent patterns of cortical thickness change across the lifespan into a small number of spatially interpretable gradients.

4.10.1 Charting Along Gradients

SNM's spatial query framework also supports the characterization of normative cortical trajectories along continuous spatial gradients. For each growth gradient (TGG1–3), we stratified the cortex at regular intervals along the gradient using a probabilistic approach. First, the vertex-wise values of each gradient map were transformed into centile maps ranging from 0 to 100 across all vertices. We defined 11 target centiles, $C = [0, 10, 20, \ldots, 100]$, to serve as reference points along each gradient.

For each centile c_j, we applied a Gaussian weighting function to the gradient map i to define smooth gradient strata:

$$\text{gradient strata}(i, j) = e^{-\left(\frac{TGG_i - c_j}{\sigma}\right)^2} \tag{28}$$

where σ denotes the kernel width and was set to 10 centiles FWHM. This probabilistic weighting allows vertices near the target centile to contribute strongly to its normative trajectory. Each stratum was normalized to sum to one, producing a spatial query that represents the weighted average cortical thickness of regions near the target centile along the gradient (see Extended Data Figs. E.2, E.3, E.4 for illustrative examples). Finally, the pretrained SNM was used to estimate normative cortical thickness trajectories of these spatial queries (Fig. 4B,C, and Extended Data Figs. E.5, E.6, E.7).

4.11 Clinical Evaluations

In Section 2.5 SNM Uncovers Cortical Signatures of Atrophy in Alzheimer's Disease and Section 2.6 SNM Highlights Heterogeneity in Individual Cortical Atrophy, we examined the clinical insights provided by SNM when applied to the clinical MACC dataset. A subset of 66 healthy individuals (50% of the HC data) from the MACC cohort was used to fine-tune the pretrained SNM for the clinical sample. During fine-tuning, parameters linking cortical phenotypes to age and sex $(w_\mu, w_\sigma^+, w_\rho, \alpha_\mu, \alpha_\sigma^+, \alpha_\rho)$ were held constant, while only site-specific parameters for batch effect harmonization $(\beta_\mu, \beta_\sigma^+, \beta_\rho)$ were updated based on the MACC data. The fine-tuned model was then applied to generate normative deviation maps for smoothed, vertex-resolution cortical thickness phenotypes.

4.11.1 HC vs. AD Comparison

A two-sample t-test was conducted to compare high-resolution deviation maps between the healthy and AD groups within the MACC dataset. To identify vertices with statistically significant differences, a nonparametric false discovery rate (FDR) correction[179] was applied across all vertices (FDR-$\alpha = 5\%$, 1000 permutations), controlling for multiple comparisons. Vertices surpassing the corrected threshold were deemed to exhibit significantly different deviation patterns between groups.

4.11.2 Linking Atrophy to Cognitive Impairment

We next evaluated the potential of normative deviation maps to serve as biomarkers of cognitive impairment. To begin, we performed a univariate vertex-wise assessment, measuring the Pearson correlation coefficient between normative z-scores and cognitive performance scores (MMSE). The significance of associations was evaluated with a nonparametric FDR correction via permutation testing. Specifically, we shuffled MMSE scores across 1000 permutations, calculating correlation coefficients for each permutation to produce a null distribution under the hypothesis of no association between normative deviation and cognitive performance. Using this null distribution, we derived nonparametric p-values for each vertex, which were subsequently FDR-corrected at $\alpha = 5\%$.

In addition, we derived an extreme value statistic[48] for atrophy by counting vertices with a z-score below a threshold of -1.96 for each individual. This measure, termed Extremely Thin Vertex Count (ETVC), was then tested for its predictive power on MMSE scores via Pearson correlation. To assess the sensitivity of ETVC to the threshold choice, we repeated this analysis for a range of z-score thresholds. We performed these assessments within each cohort (HC, MCI, and AD) to evaluate the specificity of ETVC as a biomarker for AD-related cognitive impairment, noting that no meaningful association is expected in HC, weaker associations may be present in MCI, and stronger associations are anticipated in AD (see Supplementary Section S.1.9 Within-Group Cognitive Associations).

4.11.3 Examining Interindividual Heterogeneity

Next, we aimed to demonstrate the capabilities of the SNM in revealing the heterogeneity landscape of brain atrophy in AD. We quantified interindividual differences in atrophy by computing the Euclidean distance between vectors of vertex-level z-score deviation maps, yielding a $N_{MACC} \times N_{MACC}$ distance matrix that captures interindividual differences. This matrix was visualized as a heatmap to provide

an overview of the variation across individuals. Additionally, we applied multidimensional scaling (MDS)[180] to reduce the distance matrix to a 2-dimensional embedding, providing a spatial representation of interindividual differences. Average deviation maps for subgroups within local vicinities in this 2-dimensional space were then visualized as cortical projections (see Fig. 6 and Supplementary Section S.1.12 Heterogeneity Landscape Subgroups).

To quantitatively verify the illustrated heterogeneity, we compared distance distributions across groups using two-sample t-tests and computed Cohen's d effect sizes. These comparisons test the hypothesis that individuals with MCI and AD exhibit progressively greater deviations from healthy norms, as well as increased interindividual variability within each diagnostic group, consistent with the observed heterogeneity landscape. In addition, the intrinsic cluster structure of interindividual differences was evaluated using Hartigan's dip statistic[49], which tests for multimodality in the distribution of pairwise distances. This analysis provides a quantitative assessment of whether deviation patterns exhibit evidence of cluster structure or instead vary continuously along a heterogeneous spectrum without clear boundaries[50]. Together, these analyses complement the visualizations by quantitatively characterizing both the magnitude and structural organization of heterogeneity within and between diagnostic groups.

Code and Data Availability

All scripts used to conduct the analyses and generate the figures in this study are openly available at `https://github.com/sina-mansour/normative_brain_charts`. The repository additionally hosts supplementary resources, including the pretrained SNM-1000 model, derived annual normative charts for several established brain atlases, as well as cortical maps of thickness growth gradients.

To facilitate broader adoption and reuse, the methodological components of the SNM framework are provided in the `spectranorm` Python software package (`https://sina-mansour.github.io/spectranorm/`). The accompanying website provides comprehensive documentation and tutorials describing the range of normative modeling functionalities supported by the software.

This study integrates multiple large-scale neuroimaging datasets (see Section S.1.1 Healthy Lifespan Imaging Data), each subject to its own data access and usage agreements. While some datasets are publicly available, others require approval through data access committees or the execution of material transfer or data-use agreements. Researchers interested in accessing these datasets should consult the corresponding data providers for specific access procedures.

Acknowledgments

The analysis for this study was supported by the Spartan High-Performance Computing infrastructure[181,182] and dedicated computing and storage solutions provided by Research Computing Services at the University of Melbourne. We gratefully acknowledge the invaluable contribution of several open-source software packages that significantly facilitated our data analysis and interpretation. The analytical pipeline made use of several open-source Python packages, including Numpy[183], Scipy[184], Matplotlib[185], Pandas[186], scikit-learn[187], PyMC[188], xarray[189], and Nibabel[190]. The connectome spatial smoothing package[166] was utilized to compute smoothing kernels. Brain visualizations were generated by the Cerebro brain viewer[191]. This study was supported by the NHMRC Ideas Grant (GNT2011592) titled *Normative Reference Ranges for Brain Phenotypes*, awarded to A. Zalesky, C. Bousman, C. Pantelis, and M. Di Biase; the NUS Yong Loo Lin School of Medicine (NUHSRO/2020/124/TMR/LOA); the Singapore National Medical Research Council (NMRC) OF-YIRG (OFYIRG25jan-0049), LCG (OFLCG19May-0035), NMRC CTG-IIT (CTGIIT23jan-0001), NMRC OF-IRG (OFIRG24jan-0030), NMRC STaR (STaR20nov-0003); Singapore Ministry of Health (MOH) Centre Grant (CG21APR1009); the Temasek Foundation (TF2223-IMH-01); and the United States National Institutes of Health (R01MH133334 to B.T.T.Y. and R01MH133843 to A.A.B.). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not reflect the views of the funders.

We gratefully acknowledge the use of several openly shared MRI biobank initiatives that made this work possible, including OpenNeuro (`https://openneuro.org/`), the Reproducible Brain Charts

(https://reprobrainchart.github.io/), the Healthy Brain Network (https://healthybrainnetwork.org/), UK Biobank (https://www.ukbiobank.ac.uk/), the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org/), the Human Connectome Project (http://www.humanconnectomeproject.org/), data made available by the Open Science Framework (https://osf.io/), the Laboratory of NeuroImaging (https://loni.usc.edu/), the International Neuroimaging Datasharing Initiative (INDI) (https://fcon_1000.projects.nitrc.org/), and the NIMH Data Archive (https://nda.nih.gov/).

HCP Young Adult data used in this study were provided by the Human Connectome Project, WU–Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657), funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems Neuroscience at Washington University. Data from the HCP Aging cohort used in this study were supported by the National Institute on Aging of the National Institutes of Health under Award Number U01AG052564 and by funds provided by the McDonnell Center for Systems Neuroscience at Washington University in St. Louis. The HCP-Aging 2.0 Release data used in this report were obtained from DOI: 10.15154/1520707. Data from the HCP Development cohort used in this study were supported by the National Institute of Mental Health of the National Institutes of Health under Award Number U01MH109589 and by funds provided by the McDonnell Center for Systems Neuroscience at Washington University in St. Louis. The HCP-Development 2.0 Release data used in this report were obtained from DOI: 10.15154/1520708.

Data from the Adolescent Brain Cognitive Development® (ABCD) Study used in this work were obtained from the NIH Brain Development Cohorts (NBDC) Data Sharing Platform. The ABCD Study is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9-10 years and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, and U24DA041147. ABCD Consortium investigators designed and implemented the ABCD study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD Consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from the ABCD Release 6.0 (NIMH Data Archive DOI: https://doi.org/10.82525/jy7n-g441).

UK Biobank data used in this study were accessed under Application Number 60698. Data from the Brazilian High-Risk Cohort (BHRC), the Philadelphia Neurodevelopmental Cohort (PNC), the developing Chinese Connectome Project (dCCNP), the Healthy Brain Network (HBN), and the Nathan Kline Institute (NKI) were accessed through the Reproducible Brain Charts (RBC) platform (https://reprobrainchart.github.io/). Data were provided in part by the Brain Genomics Superstruct Project (BGSP) of Harvard University and the Massachusetts General Hospital, (Principal Investigators: Randy Buckner, Joshua Roffman, and Jordan Smoller), with support from the Center for Brain Science Neuroinformatics Research Group, the Athinoula A. Martinos Centerfor Biomedical Imaging, and the Center for Human Genetic Research. 20 individual investigators at Harvard and MGH generously contributed data to the overall project.

Data used in preparation of this article were in part obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu/). The investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A full list of ADNI investigators is available at https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. Data collection and sharing for ADNI is funded by the National Institute on Aging (National Institutes of Health Grant U19AG024904). The grantee organization is the Northern California Institute for Research and Education. In the past, ADNI has also received funding from the National Institute of Biomedical Imaging and Bioengineering, the Canadian Institutes of Health Research, and private sector contributions through the Foundation for the National Institutes of Health (FNIH) including generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company;

CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics.

Data for this study were provided in part by the Australian Schizophrenia Research Bank (ASRB), accessed with approval from the ASRB Access Committee. The ASRB was supported by funding from the National Health and Medical Research Council (Enabling Grant 386500), the Pratt Foundation, Ramsay Health Care, the Viertel Charitable Foundation, and the Schizophrenia Research Institute. Data were provided in part by OASIS-3: Longitudinal Multimodal Neuroimaging Dataset: Principal Investigators: T. Benzinger, D. Marcus, J. Morris; NIH P30 AG066444, P50 AG00561, P30 NS09857781, P01 AG026276, P01 AG003991, R01 AG043434, UL1 TR000448, R01 EB009352, P30 AG066444, AW00006993.

Data used in part in this study were obtained from the Australian Imaging Biomarkers and Lifestyle flagship study of ageing (AIBL), a consortium comprising Austin Health, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Edith Cowan University, the Florey Institute for Neuroscience and Mental Health (University of Melbourne), and the National Ageing Research Institute. Numerous commercial interactions have supported data collection and analyses. In-kind support has been provided by the consortium institutions, as well as Alzheimer's Research Australia, Sir Charles Gairdner Hospital, Hollywood Private Hospital, and St Vincent's Hospital. The study has also received financial support at various stages from the Alzheimer's Drug Discovery Foundation, the Victorian Government's Operational Infrastructure Support program, Alzheimer's Research Australia, and the National Health and Medical Research Council (NHMRC). The AIBL researchers contributed data but did not participate in analysis or writing of this report. AIBL researchers are listed at https://aibl.org.au/. Data for this study were provided in part by the Growing Up in Singapore Towards Healthy Outcomes (GUSTO) study. The GUSTO dataset was supported by funding from the NRF's Human Health and Potential Domain, under the Human Potential Programme. Data used in preparation of this article were in part obtained from the SG70 study. The SG70 study was supported by Singapore National Medical Research Council [CSA-SI (MOH-000434)] and Yong Loo Lin School of Medicine, National University of Singapore.

Competing interests

S.M.L., M.A.D.B., C.C., J.H.Z, B.T.T.Y., and A.Z. have filed a patent application (US Application No. 19/247,482; status: pending) with the National University of Singapore and the University of Melbourne as applicants, covering the spectral normative modeling technology described in this article. A.A.B., R.B., and J.S. hold equity in, J.S. is a director of, and R.B. is a co-founder of Centile Bioscience. R.B. is funded by an Academy of Medical Sciences Springboard award and supported by the HDR UK Molecular to Health Records programme. The remaining authors declare no competing interests.

References

1. Marquand, A. F., Rezek, I., Buitelaar, J. & Beckmann, C. F. Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. *Biological Psychiatry* **80,** 552–561. 10.1016/j.biopsych.2015.12.023 (2016).
2. Dinga, R. *et al.* Normative modeling of neuroimaging data using generalized additive models of location scale and shape. *bioRxiv.* 10.1101/2021.06.14.448106 (2021).
3. Rutherford, S. *et al.* The normative modeling framework for computational psychiatry. *Nature Protocols* **17,** 1711–1734. 10.1038/s41596-022-00696-5 (July 2022).

4. Marquand, A. F. *et al.* Conceptualizing mental disorders as deviations from normative functioning. *Molecular Psychiatry* **24,** 1415–1424. 10.1038/s41380-019-0441-1 (2019).
5. Bethlehem, R. A. *et al.* A normative modelling approach reveals age-atypical cortical thickness in a subgroup of males with autism spectrum disorder. *Communications Biology* **3.** 10.1038/s42003-020-01212-9 (2020).
6. Bethlehem, R. A. I. *et al.* Brain charts for the human lifespan. *Nature* **604,** 525–533. 10.1038/s41586-022-04554-y (Apr. 2022).
7. Mirnezami, R., Nicholson, J. & Darzi, A. Preparing for Precision Medicine. *New England Journal of Medicine* **366,** 489–491. 10.1056/NEJMp1114866 (Feb. 2012).
8. Zabihi, M. *et al.* Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using Normative Models. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **4,** 567–578. 10.1016/j.bpsc.2018.11.013 (2019).
9. Parkes, L. *et al.* Transdiagnostic dimensions of psychopathology explain individuals' unique deviations from normative neurodevelopment in brain structure. *Translational Psychiatry* **11,** 232. 10.1038/s41398-021-01342-6 (Apr. 2021).
10. Lv, J. *et al.* Individual deviations from normative models of brain structure in a large cross-sectional schizophrenia cohort. *Molecular Psychiatry* **26,** 3512–3523. 10.1038/s41380-020-00882-5 (July 2021).
11. Habes, M. *et al.* The Brain Chart of Aging: Machine-learning analytics reveals links between brain aging, white matter disease, amyloid burden, and cognition in the iSTAGING consortium of 10,216 harmonized MR scans. *Alzheimer's and Dementia* **17,** 89–102. 10.1002/alz.12178 (2021).
12. Alexander-Bloch, A. *et al.* Copy Number Variant Risk Scores Associated with Cognition, Psychopathology, and Brain Structure in Youths in the Philadelphia Neurodevelopmental Cohort. *JAMA Psychiatry* **79,** 699–709. 10.1001/jamapsychiatry.2022.1017 (2022).
13. Holz, N. E. *et al.* A stable and replicable neural signature of lifespan adversity in the adult brain. *Nature Neuroscience* **26,** 1603–1612. 10.1038/s41593-023-01410-8 (Sept. 2023).
14. Verdi, S. *et al. Personalising Alzheimer's Disease progression using brain atrophy markers* June 2023. 10.1101/2023.06.15.23291418.
15. Kia, S. M. *et al. Hierarchical bayesian regression for multi-site normative modeling of neuroimaging data* in *Medical Image Computing and Computer Assisted Intervention–MICCAI* (2020), 699–709.
16. Fraza, C. J., Dinga, R., Beckmann, C. F. & Marquand, A. F. Warped Bayesian linear regression for normative modelling of big data. *NeuroImage* **245,** 118715. 10.1016/j.neuroimage.2021.118715 (Dec. 2021).
17. Rutherford, S. *et al.* Charting brain growth and aging at high spatial precision. *eLife* **11,** 1–15. 10.7554/eLife.72904 (Feb. 2022).
18. Wolfers, T. *et al.* Individual differences v. the average patient: Mapping the heterogeneity in ADHD using normative models. *Psychological Medicine* **50,** 314–323. 10.1017/S0033291719000084 (2019).
19. Petersen, M. *et al.* Brain network architecture constrains age-related cortical thinning. *NeuroImage* **264,** 119721. 10.1016/j.neuroimage.2022.119721 (2022).
20. Kalantar-Hormozi, H. *et al.* A cross-sectional and longitudinal study of human brain development: The integration of cortical thickness, surface area, gyrification index, and cortical curvature into a unified analytical framework. *NeuroImage* **268,** 119885. 10.1016/j.neuroimage.2023.119885 (2023).
21. Zhang, R., Chen, L., Oliver, L. D., Voineskos, A. N. & Park, J. Y. SAN: Mitigating spatial covariance heterogeneity in cortical thickness data collected from multiple scanners or sites. *Human Brain Mapping* **45,** 1–14. 10.1002/hbm.26692 (2024).

22. Pang, J. C. *et al.* Geometric constraints on human brain function. *Nature* **618,** 566–574. 10.1038/s41586-023-06098-1 (June 2023).

23. Mansour L., S. *et al.* *Eigenmodes of the brain: revisiting connectomics and geometry* Apr. 2024. 10.1101/2024.04.16.589843.

24. Olsen, A. S. *et al.* *On reconstruction of cortical functional maps using subject-specific geometric and connectome eigenmodes* Oct. 2024. 10.1101/2024.10.28.620635.

25. Robinson, P. *et al.* Eigenmodes of brain activity: Neural field theory predictions and comparison with experiment. *NeuroImage* **142,** 79–98. 10.1016/j.neuroimage.2016.04.050 (Nov. 2016).

26. Deslauriers-Gauthier, S., Zucchelli, M., Frigo, M. & Deriche, R. A unified framework for multimodal structure–function mapping based on eigenmodes. *Medical Image Analysis* **66,** 101799. 10.1016/j.media.2020.101799 (2020).

27. Atasoy, S., Donnelly, I. & Pearson, J. Human brain networks function in connectome-specific harmonic waves. *Nature Communications* **7,** 10340. 10.1038/ncomms10340 (Apr. 2016).

28. Fornito, A., Zalesky, A. & Breakspear, M. The connectomics of brain disorders. *Nature Reviews Neuroscience* **16,** 159–172. 10.1038/nrn3901 (Mar. 2015).

29. Pievani, M., Filippini, N., Van Den Heuvel, M. P., Cappa, S. F. & Frisoni, G. B. *Brain connectivity in neurodegenerative diseases - From phenotype to proteinopathy* 2014. 10.1038/nrneurol.2014.178.

30. Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L. & Greicius, M. D. Neurodegenerative Diseases Target Large-Scale Human Brain Networks. *Neuron* **62,** 42–52. 10.1016/j.neuron.2009.03.024 (2009).

31. Wannan, C. M. *et al.* Evidence for network-based cortical thickness reductions in schizophrenia. *American Journal of Psychiatry* **176,** 552–563. 10.1176/appi.ajp.2019.18040380 (2019).

32. Vogel, J. W. *et al.* Connectome-based modelling of neurodegenerative diseases: towards precision medicine and mechanistic insight. *Nature Reviews Neuroscience* **24,** 620–639. 10.1038/s41583-023-00731-8 (2023).

33. Chopra, S. *et al.* Network-Based Spreading of Gray Matter Changes Across Different Stages of Psychosis. *JAMA Psychiatry* **80,** 1246–1257. 10.1001/jamapsychiatry.2023.3293 (2023).

34. Sebenius, I. *et al.* Structural MRI of brain similarity networks. *Nature Reviews Neuroscience* **26.** 10.1038/s41583-024-00882-2 (2024).

35. Wang, X. *et al.* Profiles of Precentral and Postcentral Cortical Mean Thicknesses in Individual Subjects over Acute and Subacute Time-Scales. *Cerebral Cortex* **20,** 1513–1522. 10.1093/cercor/bhp226 (July 2010).

36. Chiarello, C., Vazquez, D., Felton, A. & McDowell, A. Structural asymmetry of the human cerebral cortex: Regional and between-subject variability of surface area, cortical thickness, and local gyrification. *Neuropsychologia* **93,** 365–379. 10.1016/j.neuropsychologia.2016.01.012 (2016).

37. Douaud, G. *et al.* A common brain network links development, aging, and vulnerability to disease. *Proceedings of the National Academy of Sciences of the United States of America* **111,** 17648–17653. 10.1073/pnas.1410378111 (2014).

38. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536,** 171–178. 10.1038/nature18933 (Aug. 2016).

39. Dear, R. *et al.* Cortical gene expression architecture links healthy neurodevelopment to the imaging, transcriptomics and genetics of autism and schizophrenia. *Nature Neuroscience* **27,** 1075–1086. 10.1038/s41593-024-01624-4 (2024).

40. Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences* **113,** 12574–12579. 10.1073/pnas.1608282113 (Nov. 2016).

41. Sydnor, V. J. *et al.* Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron* **109,** 2820–2846. 10.1016/j.neuron.2021.06.016 (2021).

42. Verdi, S. *et al.* Revealing Individual Neuroanatomical Heterogeneity in Alzheimer Disease Using Neuroanatomical Normative Modeling. *Neurology* **100,** E2442–E2453. 10.1212/WNL.0000000000207298 (June 2023).

43. Loreto, F. *et al.* Alzheimer's disease heterogeneity revealed by neuroanatomical normative modeling. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring* **16,** 1–13. 10.1002/dad2.12559 (2024).

44. Fjell, A. M., McEvoy, L., Holland, D., Dale, A. M. & Walhovd, K. B. What is normal in normal aging? Effects of aging, amyloid and Alzheimer's disease on the cerebral cortex and the hippocampus. *Progress in Neurobiology* **117,** 20–40. 10.1016/j.pneurobio.2014.02.004 (June 2014).

45. Bakkour, A., Morris, J. C. & Dickerson, B. C. The cortical signature of prodromal AD: Regional thinning predicts mild AD dementia. *Neurology* **72,** 1048–1055. 10.1212/01.wnl.0000340981.97664.2f (2009).

46. Dickerson, B. C. *et al.* The cortical signature of Alzheimer's disease: Regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. *Cerebral Cortex* **19,** 497–510. 10.1093/cercor/bhn113 (2009).

47. Poulakis, K. *et al.* Heterogeneous patterns of brain atrophy in Alzheimer's disease. *Neurobiology of Aging* **65,** 98–108. 10.1016/j.neurobiolaging.2018.01.009 (2018).

48. Fraza, C., Zabihi, M., Beckmann, C. F. & Marquand, A. F. *Reconceptualizing psychopathology as extreme deviations from a normative reference model* Aug. 2022. 10.1101/2022.08.23.505049.

49. Hartigan, P. M. Algorithm AS 217: Computation of the Dip Statistic to Test for Unimodality. *Applied Statistics* **34,** 320. 10.2307/2347485 (1985).

50. Adolfsson, A., Ackerman, M. & Brownstein, N. C. To cluster, or not to cluster: An analysis of clusterability methods. *Pattern Recognition* **88,** 13–26. 10.1016/j.patcog.2018.10.026 (2019).

51. Burt, J. B. *et al.* Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nature Neuroscience* **21,** 1251–1259. 10.1038/s41593-018-0195-0 (2018).

52. Huntenburg, J. M., Bazin, P. L. & Margulies, D. S. Large-Scale Gradients in Human Cortical Organization. *Trends in Cognitive Sciences* **22,** 21–31. 10.1016/j.tics.2017.11.002 (2018).

53. Sydnor, V. J. *et al.* Intrinsic activity development unfolds along a sensorimotor–association cortical axis in youth. *Nature Neuroscience* **26,** 638–649. 10.1038/s41593-023-01282-y (2023).

54. Fjell, A. M. *et al.* Development and aging of cortical thickness correspond to genetic organization patterns. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 15462–15467. 10.1073/pnas.1508831112 (2015).

55. Vidal-Pineiro, D. *et al.* Cellular correlates of cortical thinning throughout the lifespan. *Scientific Reports* **10,** 21803. 10.1038/s41598-020-78471-3 (Dec. 2020).

56. Shen, C., Liu, C. & Qiu, A. Metabolism-related brain morphology accelerates aging and predicts neurodegenerative diseases and stroke: a UK Biobank study. *Translational Psychiatry* **13,** 1–11. 10.1038/s41398-023-02515-1 (2023).

57. Whitaker, K. J. *et al.* Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proceedings of the National Academy of Sciences of the United States of America* **113,** 9105–9110. 10.1073/pnas.1601745113 (2016).

58. Fischl, B. Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex* **14,** 11–22. 10.1093/cercor/bhg087 (Jan. 2004).

59. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31,** 968–980. 10.1016/j.neuroimage.2006.01.021 (July 2006).

60. Destrieux, C., Fischl, B., Dale, A. & Halgren, E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage* **53,** 1–15. 10.1016/j.neuroimage.2010.06.010 (Oct. 2010).

61. Yeo, B. T. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* **106,** 1125–1165. 10.1152/jn.00338.2011 (2011).

62. Power, J. D. *et al.* Functional Network Organization of the Human Brain. *Neuron* **72,** 665–678. 10.1016/j.neuron.2011.09.006 (Nov. 2011).

63. Fan, L. *et al.* The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cerebral Cortex* **26,** 3508–3526. 10.1093/cercor/bhw157 (2016).

64. Gordon, E. M. *et al.* Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cerebral Cortex* **26,** 288–303. 10.1093/cercor/bhu239 (2016).

65. Schaefer, A. *et al.* Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex* **28,** 3095–3114. 10.1093/cercor/bhx179 (2018).

66. Eickhoff, S. B., Yeo, B. T. T. & Genon, S. Imaging-based parcellations of the human brain. *Nature Reviews Neuroscience* **19,** 672–686. 10.1038/s41583-018-0071-7 (Nov. 2018).

67. Poline, J. B. *et al.* Data sharing in neuroimaging research. *Frontiers in Neuroinformatics* **6,** 1–13. 10.3389/fninf.2012.00009 (2012).

68. Eickhoff, S., Nichols, T. E., Van Horn, J. D. & Turner, J. A. Sharing the wealth: Neuroimaging data repositories. *NeuroImage* **124,** 1065–1068. 10.1016/j.neuroimage.2015.10.079 (2016).

69. Jwa, A. S. & Poldrack, R. A. The spectrum of data sharing policies in neuroimaging data repositories. *Human Brain Mapping* **43,** 2707–2721. 10.1002/hbm.25803 (2022).

70. Giehl, K. *et al.* Sharing brain imaging data in the Open Science era: how and why? *The Lancet Digital Health* **6,** e526–e535. 10.1016/S2589-7500(24)00069-4 (2024).

71. Busche, M. A. & Hyman, B. T. Synergy between amyloid-β and tau in Alzheimer's disease. *Nature Neuroscience* **23,** 1183–1193. 10.1038/s41593-020-0687-6 (2020).

72. Hampel, H. *et al.* The Amyloid-β Pathway in Alzheimer's Disease. *Molecular Psychiatry* **26,** 5481–5503. 10.1038/s41380-021-01249-0 (2021).

73. Das, S. R. *et al.* Tau-Atrophy Variability Reveals Phenotypic Heterogeneity in Alzheimer's Disease. *Annals of Neurology* **90,** 751–762. 10.1002/ana.26233 (2021).

74. Frigerio, I. *et al.* Amyloid-β, p-tau and reactive microglia are pathological correlates of MRI cortical atrophy in Alzheimer's disease. *Brain Communications* **3,** 1–13. 10.1093/braincomms/fcab281 (2021).

75. Scahill, R. I., Schott, J. M., Stevens, J. M., Rossor, M. N. & Fox, N. C. Mapping the evolution of regional atrophy in Alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. *Proceedings of the National Academy of Sciences of the United States of America* **99,** 4703–4707. 10.1073/pnas.052587399 (2002).

76. Dickerson, B. C. *et al.* Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. *Neurology* **76,** 1395–1402. 10.1212/WNL.0b013e3182166e96 (2011).

77. Pan, S. J. & Yang, Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering* **22,** 1345–1359. 10.1109/TKDE.2009.191 (2010).

78. Dufumier, B. *et al.* Exploring the potential of representation and transfer learning for anatomical neuroimaging: Application to psychiatry. *NeuroImage* **296,** 120665. 10.1016/j.neuroimage.2024.120665 (2024).

79. Zetterberg, H. & Bendlin, B. B. Biomarkers for Alzheimer's disease—preparing for a new era of disease-modifying therapies. *Molecular Psychiatry* **26,** 296–308. 10.1038/s41380-020-0721-9 (2021).

80. Whitwell, J. L. Progression of atrophy in Alzheimer's disease and related disorders. *Neurotoxicity Research* **18,** 339–346. 10.1007/s12640-010-9175-1 (2010).

81. Sun, N., Mormino, E. C., Chen, J., Sabuncu, M. R. & Yeo, B. T. Multi-modal latent factor exploration of atrophy, cognitive and tau heterogeneity in Alzheimer's disease. *NeuroImage* **201.** 10.1016/j.neuroimage.2019.116043 (2019).

82. Poulakis, K. *et al.* Multi-cohort and longitudinal Bayesian clustering study of stage and subtype in Alzheimer's disease. *Nature Communications* **13.** 10.1038/s41467-022-32202-6 (2022).

83. Jones, B. F. *et al.* Differential regional atrophy of the cingulate gyrus in Alzheimer disease: A volumetric MRI study. *Cerebral Cortex* **16,** 1701–1708. 10.1093/cercor/bhj105 (2006).

84. Benzinger, T. L. *et al.* Regional variability of imaging biomarkers in autosomal dominant Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America* **110.** 10.1073/pnas.1317918110 (2013).

85. Phan, T. X. *et al.* Increased Cortical Thickness in Alzheimer's Disease. *Annals of Neurology* **95,** 929–940. 10.1002/ana.26894 (2024).

86. Iacono, D. *et al.* Neuronal hypertrophy in asymptomatic Alzheimer disease. *Journal of Neuropathology and Experimental Neurology* **67,** 578–589. 10.1097/NEN.0b013e3181772794 (2008).

87. Xia, X., Jiang, Q., McDermott, J. & Han, J.-D. J. Aging and Alzheimer's disease: Comparison and associations from molecular to system level. *Aging Cell* **17,** 1–14. 10.1111/acel.12802 (Oct. 2018).

88. Kubben, N. & Misteli, T. *Shared molecular and cellular mechanisms of premature ageing and ageing-associated diseases* Oct. 2017. 10.1038/nrm.2017.68.

89. Verdi, S., Marquand, A. F., Schott, J. M. & Cole, J. H. Beyond the average patient: How neuroimaging models can address heterogeneity in dementia. *Brain* **144,** 2946–2953. 10.1093/brain/awab165 (2021).

90. Ownby, R. L., Crocco, E., Acevedo, A., John, V. & Loewenstein, D. Depression and Risk for Alzheimer Disease. *Archives of General Psychiatry* **63,** 530. 10.1001/archpsyc.63.5.530 (2006).

91. Körner, S. *et al.* Prevalence and prognostic impact of comorbidities in amyotrophic lateral sclerosis. *European Journal of Neurology* **20,** 647–654. 10.1111/ene.12015 (2013).

92. Wang, J. H., Wu, Y. J., Tee, B. L. & Lo, R. Y. Medical Comorbidity in Alzheimer's Disease: A Nested Case-Control Study. *Journal of Alzheimer's Disease* **63,** 773–781. 10.3233/JAD-170786 (2018).

93. Avitan, I. *et al.* Towards a consensus on alzheimer's disease comorbidity? *Journal of Clinical Medicine* **10,** 1–13. 10.3390/jcm10194360 (2021).

94. Whitwell, J. L. *et al.* Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: A case-control study. *The Lancet Neurology* **11,** 868–877. 10.1016/S1474-4422(12)70200-4 (2012).

95. Noh, Y. *et al.* Anatomical heterogeneity of Alzheimer disease. *Neurology* **83,** 1936–1944. 10.1212/WNL.0000000000001003 (Nov. 2014).

96. Zhang, X. *et al.* Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America* **113,** E6535–E6544. 10.1073/pnas.1611073113 (2016).

97. Ferreira, D. *et al.* Distinct subtypes of Alzheimer's disease based on patterns of brain atrophy: Longitudinal trajectories and clinical applications. *Scientific Reports* **7,** 1–13. 10.1038/srep46263 (2017).

98. Vogel, J. W. *et al.* Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nature Medicine* **27,** 871–881. 10.1038/s41591-021-01309-6 (May 2021).

99. Verdi, S. *et al.* Revealing Individual Neuroanatomical Heterogeneity in Alzheimer Disease Using Neuroanatomical Normative Modeling. *Neurology* **100,** E2442–E2453. 10.1212/WNL.0000000000207298 (2023).

100. Bhome, R. *et al.* A neuroimaging measure to capture heterogeneous patterns of atrophy in Parkinson's disease and dementia with Lewy bodies. *NeuroImage: Clinical* **42,** 103596. 10.1016/j.nicl.2024.103596 (2024).

101. Mohanty, R. *et al.* Comparison of subtyping methods for neuroimaging studies in Alzheimer's disease: A call for harmonization. *Brain Communications* **2,** 1–16. 10.1093/braincomms/fcaa192 (2020).

102. Hohman, T. J. *et al.* Asymptomatic Alzheimer disease. *Neurology* **87,** 2443–2450. 10.1212/WNL.0000000000003397 (2016).

103. Gómez-Isla, T. & Frosch, M. P. Lesions without symptoms: understanding resilience to Alzheimer disease neuropathological changes. *Nature Reviews Neurology* **18,** 323–332. 10.1038/s41582-022-00642-9 (2022).

104. Yang, Z. *et al.* Brain aging patterns in a large and diverse cohort of 49,482 individuals. *Nature Medicine* **30,** 3015–3026. 10.1038/s41591-024-03144-x (Oct. 2024).

105. Chouliaras, L. & O'Brien, J. T. The use of neuroimaging techniques in the early and differential diagnosis of dementia. *Molecular Psychiatry* **28,** 4084–4097. 10.1038/s41380-023-02215-8 (2023).

106. Huang, J. *et al.* Cerebral microinfarcts revisited: Detection, causes, and clinical relevance. *International Journal of Stroke* **19,** 7–15. 10.1177/17474930231187979 (2024).

107. Pinaya, W. H. *et al.* Using normative modelling to detect disease progression in mild cognitive impairment and Alzheimer's disease in a cross-sectional multi-cohort study. *Scientific Reports* **11,** 1–13. 10.1038/s41598-021-95098-0 (2021).

108. Kumar, S., Payne, P. & Sotiras, A. Normative Modeling using Multimodal Variational Autoencoders to Identify Abnormal Brain Structural Patterns in Alzheimer Disease, 1–10 (2021).

109. Wang, X. *et al.* Normative Modeling Via Conditional Variational Autoencoder and Adversarial Learning to Identify Brain Dysfunction in Alzheimer's Disease. *Proceedings - International Symposium on Biomedical Imaging* **2023-April,** 1–4. 10.1109/ISBI53787.2023.10230377 (2023).

110. Jaramillo, C., Trujillo, M. & Bernal, J. Deep learning based normative models identify structures that become abnormal with Alzheimer's disease. *Alzheimer's & Dementia* **19,** 1–4. 10.1002/alz.075865 (2023).

111. Lawry Aguila, A., Lorenzini, L., Janahi, M., Barkhof, F. & Altmann, A. Deep normative modelling reveals insights into early-stage Alzheimer's disease using multi-modal neuroimaging data. *Alzheimer's Research and Therapy* **17.** 10.1186/s13195-025-01753-3 (2025).

112. Tarun, A., Abramian, D., Behjat, H. & De Ville, D. V. *Graph Spectral Analysis of Voxel-Wise Brain Graphs from Diffusion-Weighted Mri* in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (IEEE, Apr. 2019), 159–163. 10.1109/ISBI.2019.8759496.

113. Villalón-Reina, J. E. *et al.* *Large-scale Normative Modeling of Brain Microstructure* in *2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM)* (IEEE, Nov. 2023), 1–5. 10.1109/SIPAIM56729.2023.10373451.

114. De Boer, A. A. A. *et al.* Non-Gaussian normative modelling with hierarchical Bayesian regression. *Imaging Neuroscience* **2,** 1–36. 10.1162/imag{_}a{_}00132 (Apr. 2024).

115. Rutherford, S. *et al.* Evidence for embracing normative modeling. *eLife* **12,** 1–24. 10.7554/eLife.85082 (Mar. 2023).

116. Kasper, J. *et al.* Resting state changes in aging and Parkinson's disease are shaped by underlying neurotransmission – a normative modeling study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging,* 1–12. 10.1016/j.bpsc.2024.04.010 (2024).

117. Therriault, J. *et al.* Biomarker-based staging of Alzheimer disease: rationale and clinical applications. *Nature Reviews Neurology* **20,** 232–244. 10.1038/s41582-024-00942-2 (2024).

118. Pang, J. C. *et al.* Reply to: Commentary on Pang et al. (2023). *Nature,* 1–7 (2023).

119. Becker, C. O. *et al.* Spectral mapping of brain functional connectivity from diffusion imaging. *Scientific Reports* **8,** 1–15. 10.1038/s41598-017-18769-x (2018).

120. Gavili, A. & Zhang, X.-P. On the Shift Operator, Graph Frequency, and Optimal Filtering in Graph Signal Processing. *IEEE Transactions on Signal Processing* **65,** 6303–6318. 10.1109/TSP.2017.2752689 (Dec. 2017).

121. Benzi, M., Bertaccini, D., Durastante, F. & Simunec, I. Non-local network dynamics via fractional graph Laplacians. *Journal of Complex Networks* **8,** 1–29. 10.1093/comnet/cnaa017 (2020).

122. Casey, B. J. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience* **32,** 43–54. 10.1016/j.dcn.2018.03.001 (2018).

123. Di Martino, A. *et al.* The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry* **19,** 659–667. 10.1038/mp.2013.78 (2014).

124. Di Martino, A. *et al.* Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Scientific Data* **4,** 1–15. 10.1038/sdata.2017.10 (2017).

125. Milham, P. M., Damien, F., Maarten, M. & Stewart, H. M. The ADHD-200 Consortium: A model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in Systems Neuroscience* **6,** 1–5. 10.3389/fnsys.2012.00062 (2012).

126. Jack, C. R. *et al.* The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging* **27,** 685–691. 10.1002/jmri.21049 (2008).

127. Petersen, R. C. *et al.* Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. *Neurology* **74,** 201–209. 10.1212/WNL.0b013e3181cb3e25 (2010).

128. Ellis, K. A. *et al.* The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *International Psychogeriatrics* **21,** 672–687. 10.1017/S1041610209009405 (2009).

129. Fowler, C. *et al.* Fifteen Years of the Australian Imaging, Biomarkers and Lifestyle (AIBL) Study: Progress and Observations from 2,359 Older Adults Spanning the Spectrum from Cognitive Normality to Alzheimer's Disease. *Journal of Alzheimer's Disease Reports* **5,** 443–468. 10.3233/ADR-210005 (2021).

130. Snoek, L. *et al.* The Amsterdam Open MRI Collection, a set of multimodal MRI datasets for individual difference analyses. *Scientific Data* **8,** 85. 10.1038/s41597-021-00870-6 (Mar. 2021).

131. Loughland, C. *et al.* Australian Schizophrenia Research Bank: A database of comprehensive clinical, endophenotypic and genetic data for aetiological studies of schizophrenia. *Australian and New Zealand Journal of Psychiatry* **44,** 1029–1035. 10.3109/00048674.2010.501758 (2010).

132. Holmes, A. J. *et al.* Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Scientific Data* **2,** 1–16. 10.1038/sdata.2015.31 (2015).

133. Salum, G. A. *et al.* High risk cohort study for psychiatric disorders in childhood: rationale, design, methods and preliminary results. *International Journal of Methods in Psychiatric Research* **24,** 58–73. 10.1002/mpr.1459 (Mar. 2015).

134. Liu, S. *et al.* Chinese Color Nest Project : An accelerated longitudinal brain-mind cohort. *Developmental Cognitive Neuroscience* **52,** 101020. 10.1016/j.dcn.2021.101020 (Dec. 2021).

135. Fan, X. R. *et al.* A longitudinal resource for population neuroscience of school-age children and adolescents in China. *Scientific Data* **10,** 545. 10.1038/s41597-023-02377-8 (Aug. 2023).

136. Zuo, X. N. *et al.* An open science resource for establishing reliability and reproducibility in functional connectomics. *Scientific Data* **1,** 1–13. 10.1038/sdata.2014.49 (2014).

137. Park, D. C. *et al.* The Dallas Lifespan Brain Study: A Comprehensive Adult Lifespan Data Set of Brain and Cognitive Aging. *Scientific Data* **12,** 1–14. 10.1038/s41597-025-04847-7 (2025).

138. Soh, S. E. *et al.* Cohort profile: Growing up in Singapore towards healthy outcomes (GUSTO) birth cohort study. *International Journal of Epidemiology* **43,** 1401–1409. 10.1093/ije/dyt125 (2014).

139. Alexander, L. M. *et al.* An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific Data* **4,** 170181. 10.1038/sdata.2017.181 (Dec. 2017).

140. Bookheimer, S. Y. *et al.* The Lifespan Human Connectome Project in Aging: An overview. *NeuroImage* **185,** 335–348. 10.1016/j.neuroimage.2018.10.009 (2019).

141. Somerville, L. H. *et al.* The Lifespan Human Connectome Project in Development: A large-scale study of brain connectivity development in 5–21 year olds. *NeuroImage* **183,** 456–468. 10.1016/j.neuroimage.2018.08.050 (2018).

142. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: An overview. *NeuroImage* **80,** 62–79. 10.1016/j.neuroimage.2013.05.041 (2013).

143. Tobe, R. H. *et al.* A longitudinal resource for studying connectome development and its psychiatric associations during childhood. *Scientific Data* **9,** 300. 10.1038/s41597-022-01329-y (June 2022).

144. LaMontagne, P. J. *et al. OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease* Dec. 2019. 10.1101/2019.12.13.19014902.

145. Satterthwaite, T. D. *et al.* Neuroimaging of the Philadelphia Neurodevelopmental Cohort. *NeuroImage* **86,** 544–553. 10.1016/j.neuroimage.2013.07.064 (Feb. 2014).

146. Satterthwaite, T. D. *et al.* The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage* **124,** 1115–1119. 10.1016/j.neuroimage.2013.05.012 (Jan. 2016).

147. Wei, D. *et al.* Structural and functional brain scans from the cross-sectional Southwest University adult lifespan dataset. *Scientific Data* **5,** 180134. 10.1038/sdata.2018.134 (July 2018).

148. Lai, J. S., Chua, K. Y., Li, H. & Koh, W.-P. Diet quality from mid to late life and its association with physical frailty in late life in a cohort of Chinese adults. *Nutrition Journal* **23,** 57. 10.1186/s12937-024-00964-y (June 2024).

149. Xu, X. *et al.* The SINgapore GERiatric Intervention Study to Reduce Cognitive Decline and Physical Frailty (SINGER): Study Design and Protocol. *The Journal of Prevention of Alzheimer's Disease* **9,** 40–48. 10.14283/jpad.2022.5 (Jan. 2022).

150. Liu, W. *et al.* Longitudinal test-retest neuroimaging data from healthy young adults in southwest China. *Scientific Data* **4,** 170017. 10.1038/sdata.2017.17 (Feb. 2017).

151. Loo, E. X. L. *et al.* Cohort profile: Singapore Preconception Study of Long-Term Maternal and Child Outcomes (S-PRESTO). *European Journal of Epidemiology* **36,** 129–142. 10.1007/s10654-020-00697-2 (Jan. 2021).

152. Chopra, S. *et al.* The Transdiagnostic Connectome Project: an open dataset for studying brain-behavior relationships in psychiatry. *Scientific Data* **12,** 1–22. 10.1038/s41597-025-04895-z (2025).

153. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature Neuroscience* **19,** 1523–1536. 10.1038/nn.4393 (Nov. 2016).

154. 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. 10.1016/j.neuroimage.2017.10.034 (Feb. 2018).

155. Daugherty, A. M. & Raz, N. A virtual water maze revisited: Two-year changes in navigation performance and their neural correlates in healthy adults. *NeuroImage* **146,** 492–506. 10.1016/j.neuroimage.2016.09.044 (Feb. 2017).

156. Shafiei, G. *et al.* Reproducible Brain Charts: An open data resource for mapping brain development and its associations with mental health. *Neuron* **113,** 3758–3779. 10.1016/j.neuron.2025.08.026 (Nov. 2025).

157. Hilal, S. *et al.* Markers of Cardiac Dysfunction in Cognitive Impairment and Dementia. *Medicine* **94,** e297. 10.1097/MD.0000000000000297 (Jan. 2015).

158. Hilal, S. *et al.* Cortical cerebral microinfarcts predict cognitive decline in memory clinic patients. *Journal of Cerebral Blood Flow and Metabolism* **40,** 44–53. 10.1177/0271678X19835565 (2020).

159. Fischl, B. FreeSurfer. *NeuroImage* **62,** 774–781. 10.1016/j.neuroimage.2012.01.021 (Aug. 2012).

160. Sechidis, K., Tsoumakas, G. & Vlahavas, I. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* PART 3, 145–158 (2011). 10.1007/978-3-642-23808-6{_}10.

161. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80,** 105–124. 10.1016/j.neuroimage.2013.04.127 (Oct. 2013).

162. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical Surface-Based Analysis. **194,** 179–194 (1999).

163. Bayer, J. M. *et al.* Accommodating site variation in neuroimaging data using normative and hierarchical Bayesian models. *NeuroImage* **264,** 119699. 10.1016/j.neuroimage.2022.119699 (Dec. 2022).

164. Mansour L, S., Tian, Y., Yeo, B. T., Cropley, V. & Zalesky, A. High-resolution connectomic fingerprints: Mapping neural identity and behavior. *NeuroImage* **229,** 117695. 10.1016/j.neuroimage.2020.117695 (Apr. 2021).

165. Mansour L., S., Di Biase, M. A., Smith, R. E., Zalesky, A. & Seguin, C. Connectomes for 40,000 UK Biobank participants: A multi-modal, multi-scale brain network resource. *NeuroImage* **283,** 120407. 10.1016/j.neuroimage.2023.120407 (Dec. 2023).

166. Mansour L., S., Seguin, C., Smith, R. E. & Zalesky, A. Connectome Spatial Smoothing v.0.1.1 (v.0.1.1). *Zenodo.* 10.5281/zenodo.5746619 (2021).

167. Mansour L, S., Seguin, C., Smith, R. E. & Zalesky, A. Connectome spatial smoothing (CSS): Concepts, methods, and evaluation. *NeuroImage* **250,** 118930. 10.1016/j.neuroimage.2022.118930 (Apr. 2022).

168. Ortega, A., Frossard, P., Kovacevic, J., Moura, J. M. F. & Vandergheynst, P. Graph Signal Processing: Overview, Challenges, and Applications. *Proceedings of the IEEE* **106,** 808–828. 10.1109/JPROC.2018.2820126 (May 2018).

169. Huang, W. *et al.* A Graph Signal Processing Perspective on Functional Brain Imaging. *Proceedings of the IEEE* **106,** 868–885. 10.1109/JPROC.2018.2798928 (May 2018).

170. Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A. & Vandergheynst, P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine* **30,** 83–98. 10.1109/MSP.2012.2235192 (May 2013).

171. Sandryhaila, A. & Moura, J. M. Big Data Analysis with Signal Processing on Graphs: Representation and processing of massive data sets with irregular structure. *IEEE Signal Processing Magazine* **31,** 80–90. 10.1109/MSP.2014.2329213 (Sept. 2014).

172. Boukrab, R. & Pagès-Zamora, A. Random-walk Laplacian for frequency analysis in periodic graphs. *Sensors (Switzerland)* **21,** 1–13. 10.3390/s21041275 (2021).

173. Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V. & Bastiani, F. D. *Flexible Regression and Smoothing* 10.1201/b21973 (Chapman and Hall/CRC, Apr. 2017).

174. Huertas, I. *et al.* A Bayesian spatial model for neuroimaging data based on biologically informed basis functions. *NeuroImage* **161,** 134–148. 10.1016/j.neuroimage.2017.08.009 (2017).

175. Yan, X. *et al.* Homotopic local-global parcellation of the human cerebral cortex from resting-state functional connectivity. *NeuroImage* **273,** 120010. 10.1016/j.neuroimage.2023.120010 (2023).

176. Hammond, D. K., Vandergheynst, P. & Gribonval, R. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis* **30,** 129–150. 10.1016/j.acha.2010.04.005 (2011).

177. Rasmussen, C. E. & Williams, C. K. I. *Gaussian Processes for Machine Learning* (2006).

178. Diks, C., Panchenko, V. & Van Dijk, D. Likelihood-based scoring rules for comparing density forecasts in tails. *Journal of Econometrics* **163,** 215–230. 10.1016/j.jeconom.2011.04.001 (2011).

179. Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. Permutation inference for the general linear model. *NeuroImage* **92,** 381–397. 10.1016/j.neuroimage.2014.01.060 (2014).

180. Borg, I. & Groenen, P. *Modern Multidimensional Scaling: Theory and Applications* 10.2307/2669710 (2007).

181. Meade, B., Lafayette, L., Sauter, G. & Tosello, D. *Spartan HPC-Cloud Hybrid: Delivering Performance and Flexibility* 2017. 10.4225/49/58ead90dceaaa.

182. Lafayette, L. & Wiebelt, B. Spartan and NEMO: Two HPC-cloud hybrid implementations. *Proceedings - 13th IEEE International Conference on eScience, eScience 2017,* 458–459. 10.1109/eScience.2017.70 (2017).

183. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585,** 357–362. 10.1038/s41586-020-2649-2 (2020).

184. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **17,** 261–272. 10.1038/s41592-019-0686-2 (2020).

185. Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science \& Engineering* **9,** 90–95. 10.1109/MCSE.2007.55 (2007).

186. Pandas development team, T. *pandas-dev/pandas: Pandas* Apr. 2024. 10.5281/zenodo.10957263.

187. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12,** 2825–2830 (2011).

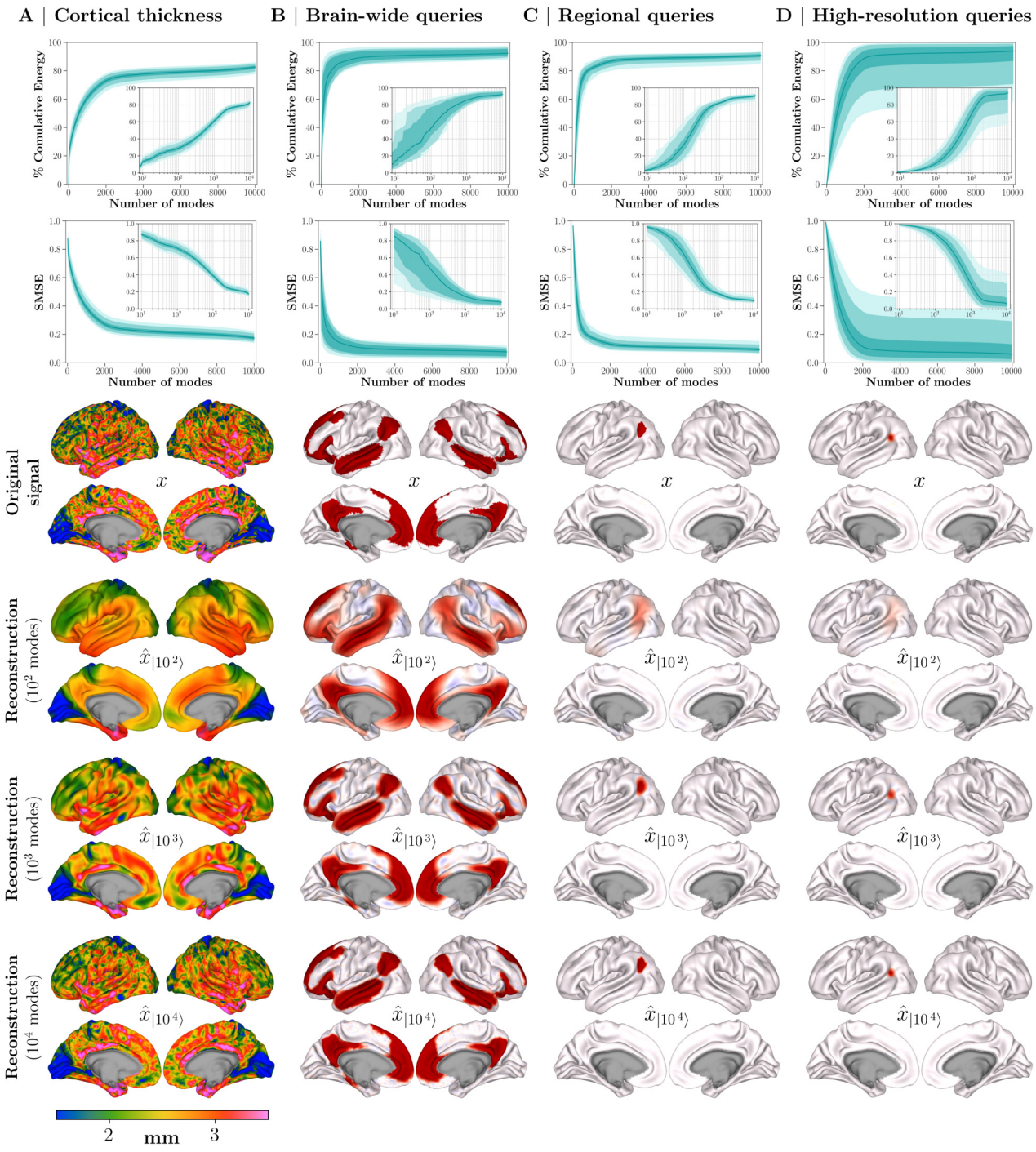
188. Abril-Pla, O. *et al.* PyMC: a modern, and comprehensive probabilistic programming framework in Python. *PeerJ Computer Science* **9,** e1516. 10.7717/peerj-cs.1516 (Sept. 2023).

189. Hoyer, S. & Hamman, J. xarray: N-D labeled Arrays and Datasets in Python. *Journal of Open Research Software* **5,** 10. 10.5334/jors.148 (2017).

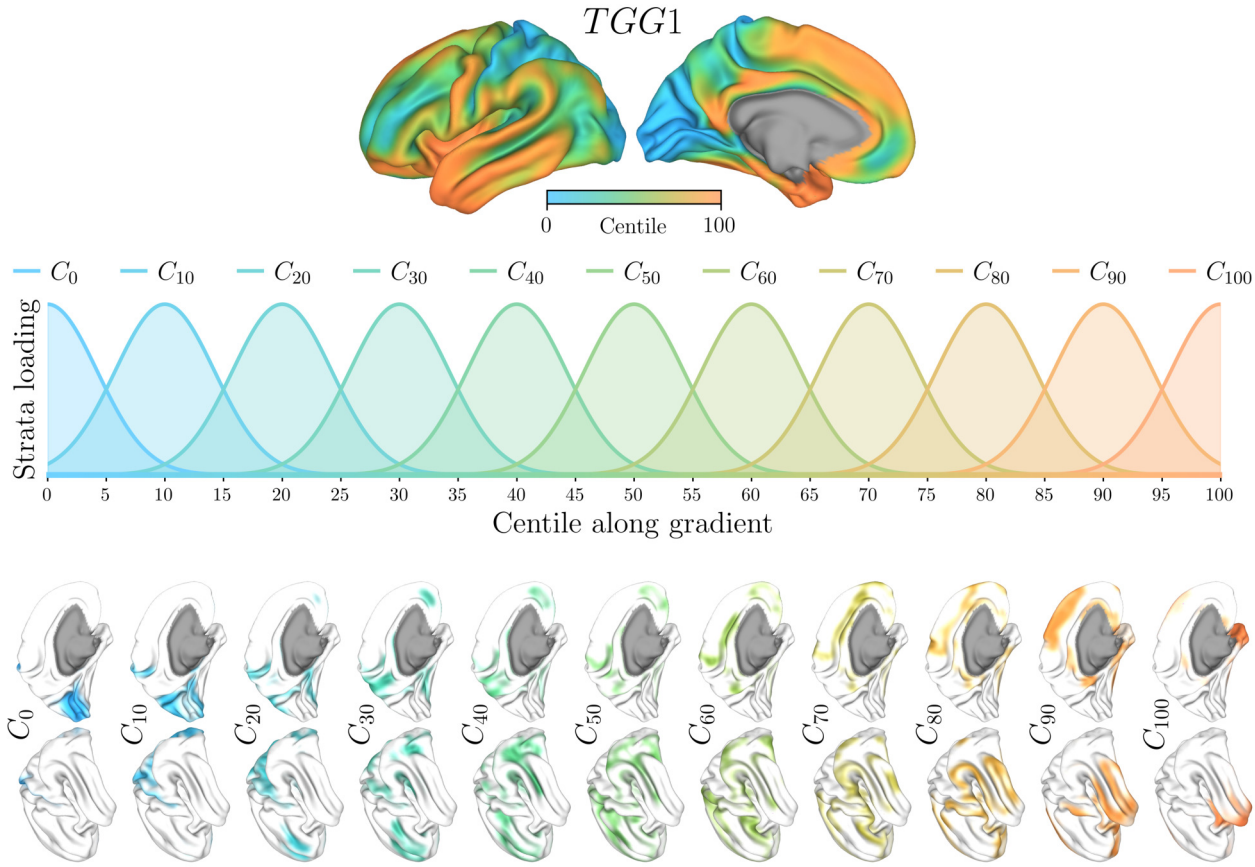
190. Brett, M. *et al. nipy/nibabel: 5.2.1* Feb. 2024. 10.5281/zenodo.10714563.

191. Mansour L., S. *et al. sina-mansour/Cerebro_Viewer: v0.0.10.4* Aug. 2023. 10.5281/zenodo.8238773.

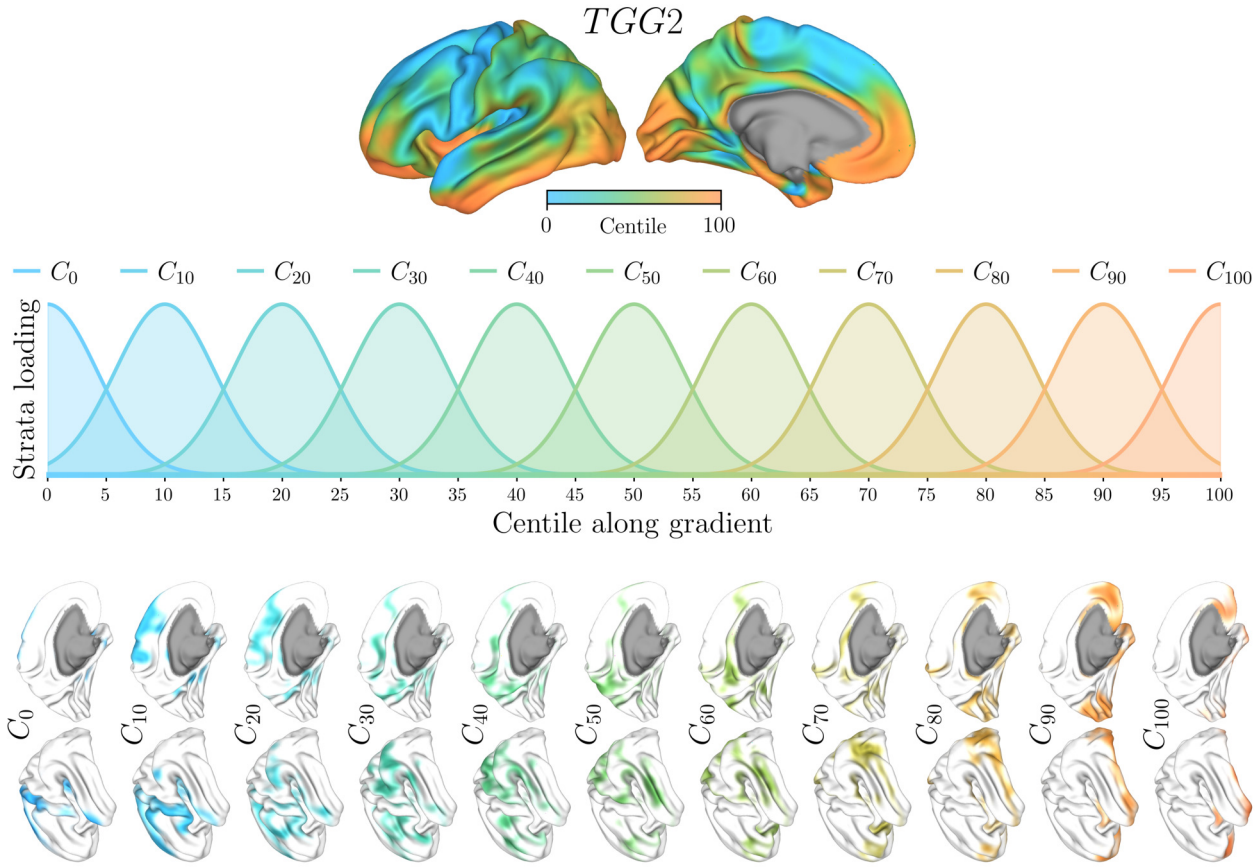
Extended Data Figures



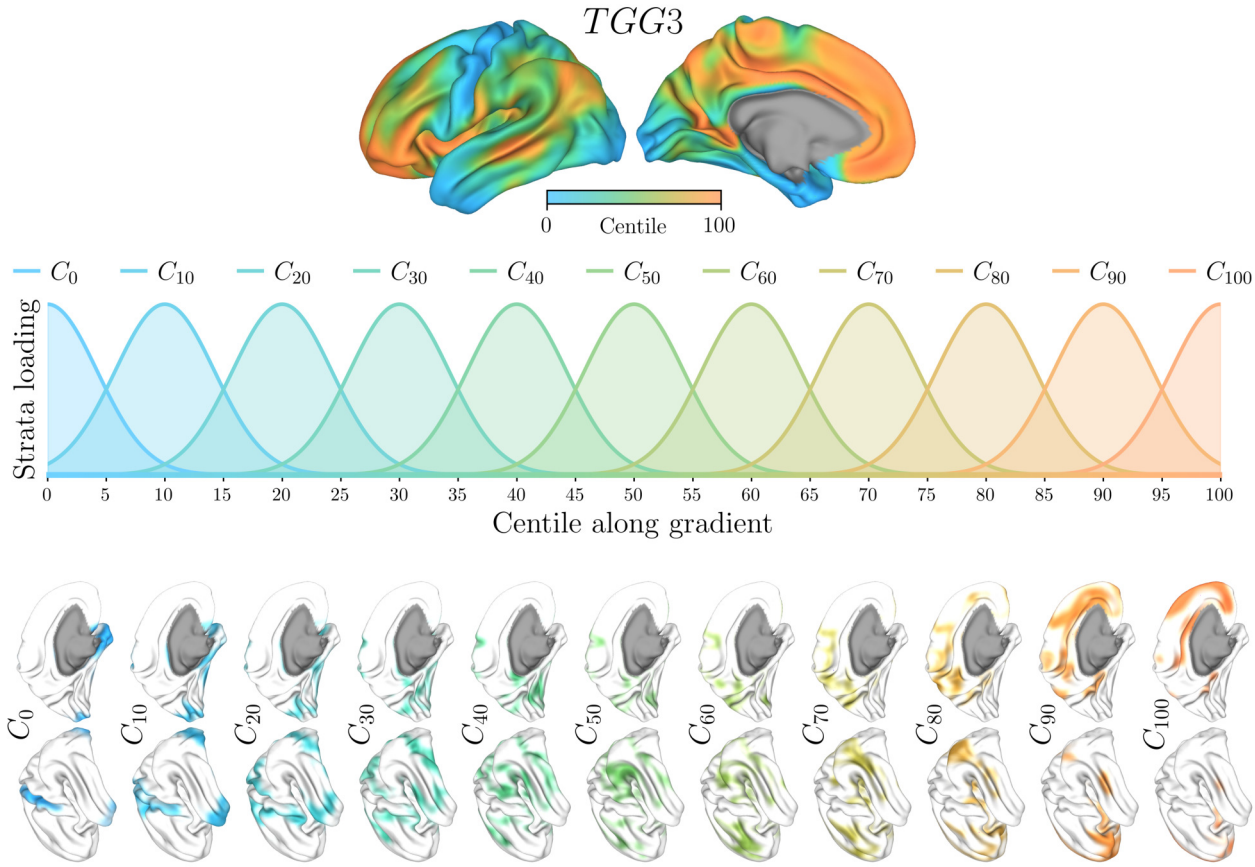
Extended Data Fig. E.1. Eigenmode-Based Reconstruction of Cortical Thickness and Spatial Query Maps. Eigenmode reconstruction accuracy for **(A)** individual cortical thickness maps and three families of spatial queries, including **(B)** brain-wide, **(C)** regional, and **(D)** high-resolution queries. Shaded line plots show the median reconstruction performance across observations, with shaded regions indicating the [25, 75], [5, 95], and [1, 99] percentiles. Cortical thickness results are based on 5,000 randomly selected individual healthy scans, while spatial query results include 25 brain-wide, 200 regional, and 4000 high-resolution queries (see Section 4.9 Model Evaluation). The first and second rows report cumulative signal energy and standardized mean squared error (SMSE), respectively, as a function of the number of low-frequency eigenmodes used for reconstruction (logarithmic x-axis shown in the insets). The third row displays an example map from each category, and subsequent rows show reconstructions using 100, 1,000, and 10,000 eigenmodes. Cortical thickness projections in panel A correspond to a single representative individual. Additional supplementary analyses examine regional variability in reconstruction accuracy, as well as the effects of spatial granularity and cortical symmetry (see Supplementary Figs. S.5, S.10, and S.13).



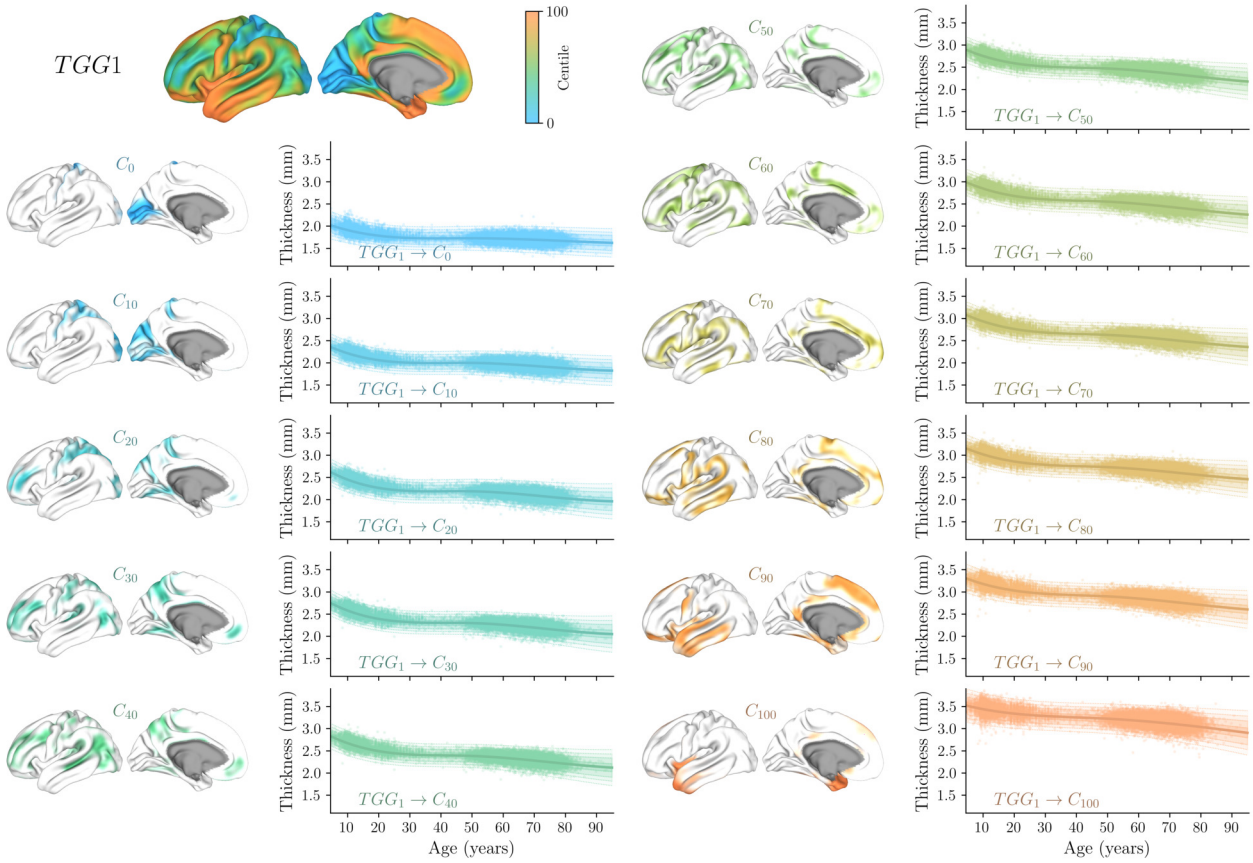
Extended Data Fig. E.2. Construction of Probabilistic Gradient Strata along TGG1. The cortical surface is stratified along the first growth gradient (TGG1) into 11 strata, each defined by a target centile, $C = [0, 10, 20, \ldots, 100]$, using probabilistic spatial query. **(A)** Vertex-wise TGG1 values transformed to centile space (0–100). **(B)** Eleven target centiles used as reference points along the gradient, with Gaussian-weighted functions defining smooth, overlapping strata centered at each centile. **(C)** Cortical projections illustrating the spatial delineation of each stratum, which are used to extract weighted mean cortical thickness estimates along the gradient.



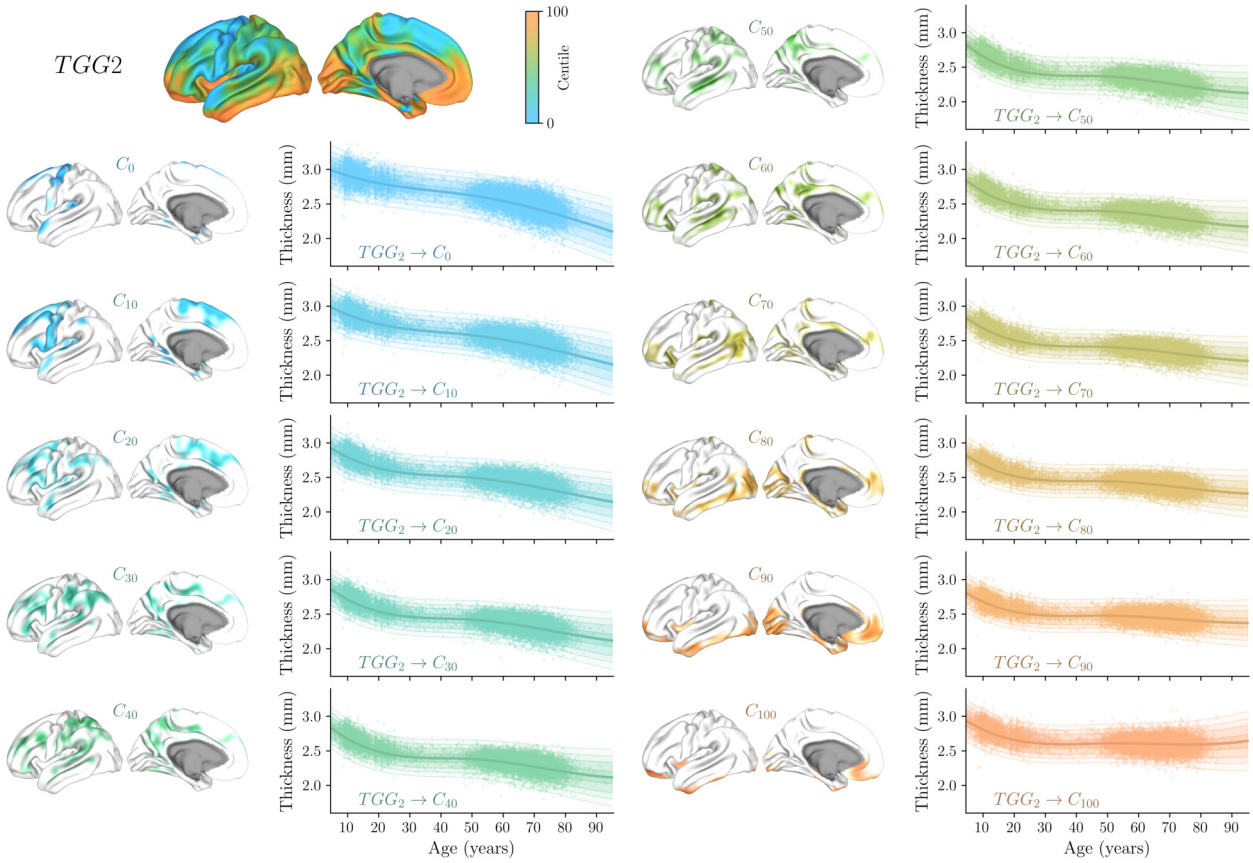
Extended Data Fig. E.3. Construction of Probabilistic Gradient Strata along TGG2. The cortical surface is stratified along the first growth gradient (TGG2) into 11 strata, each defined by a target centile, $C = [0, 10, 20, \ldots, 100]$, using probabilistic spatial query. **(A)** Vertex-wise TGG2 values transformed to centile space (0–100). **(B)** Eleven target centiles used as reference points along the gradient, with Gaussian-weighted functions defining smooth, overlapping strata centered at each centile. **(C)** Cortical projections illustrating the spatial delineation of each stratum, which are used to extract weighted mean cortical thickness estimates along the gradient.



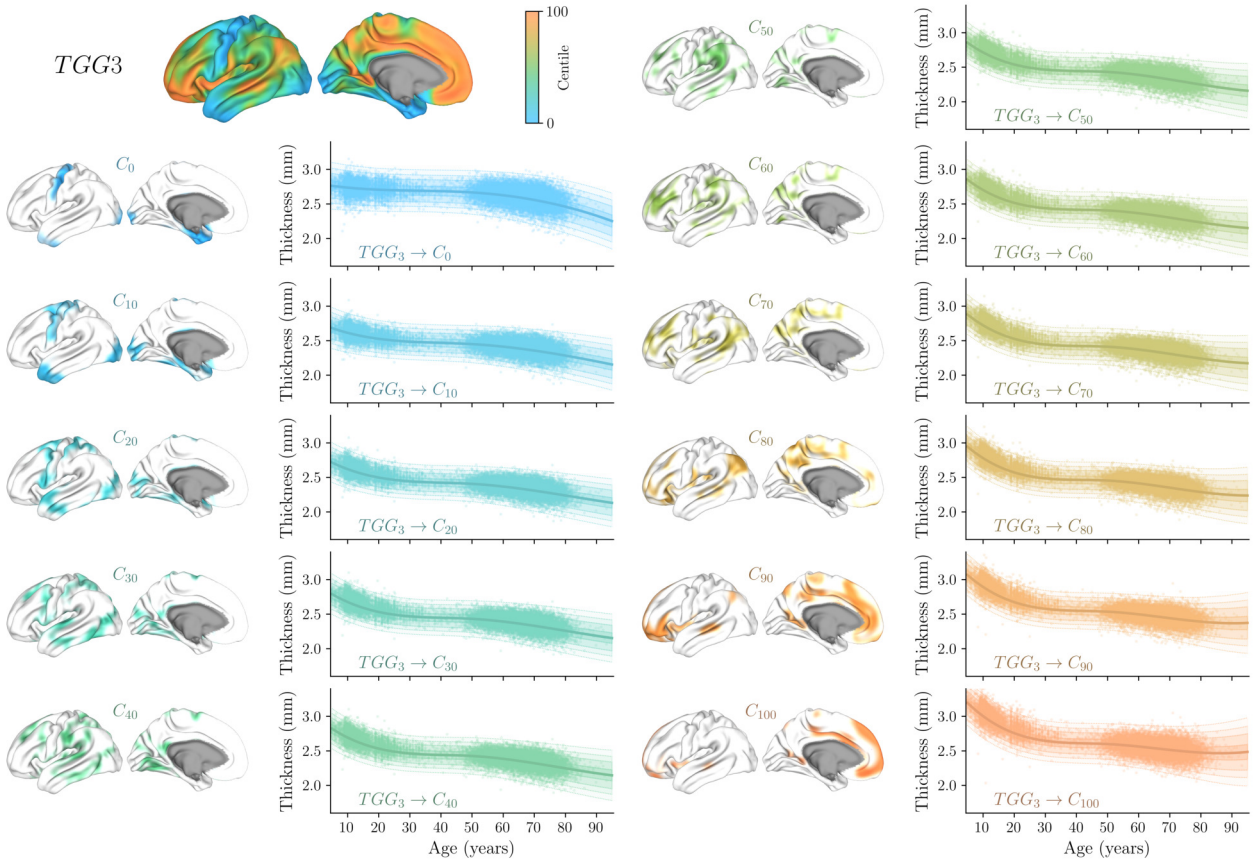
Extended Data Fig. E.4. Construction of Probabilistic Gradient Strata along TGG3. The cortical surface is stratified along the first growth gradient (TGG3) into 11 strata, each defined by a target centile, $C = [0, 10, 20, \ldots, 100]$, using probabilistic spatial query. **(A)** Vertex-wise TGG3 values transformed to centile space (0–100). **(B)** Eleven target centiles used as reference points along the gradient, with Gaussian-weighted functions defining smooth, overlapping strata centered at each centile. **(C)** Cortical projections illustrating the spatial delineation of each stratum, which are used to extract weighted mean cortical thickness estimates along the gradient.



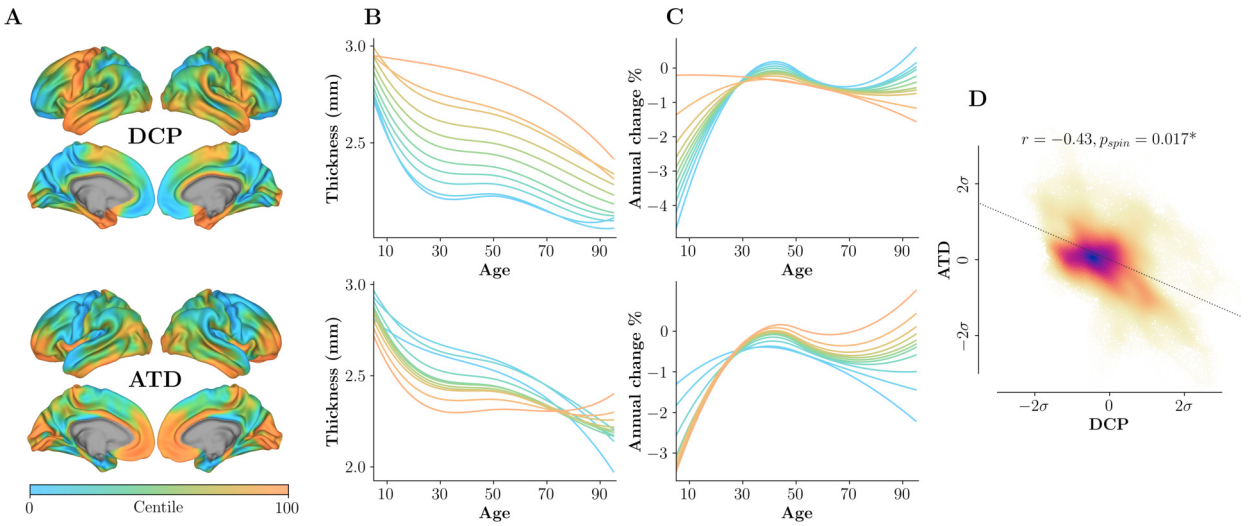
Extended Data Fig. E.5. Normative Cortical Thickness Trajectories along TGG1. Normative charts derived from spatial queries defined by probabilistic strata along the first growth gradient (TGG1, shown at the top left). For each of the eleven strata, the corresponding cortical delineation and age-dependent normative trajectory are displayed. Charts depict cortical thickness trajectories estimated using SNM-1000 as a function of age, after accounting for site and sex effects. Solid lines indicate the median estimate, with shaded regions showing the [25, 75], [5, 95], and [1, 99] percentiles. Scatter points represent queried cortical thickness values from held-out test subjects not used for model training. Together, these trajectories illustrate a stable lifelong pattern along TGG1, in which regions that are relatively thicker remain thicker and regions that are thinner remain thinner, with all strata following similar age-related profiles but differing in baseline thickness.



Extended Data Fig. E.6. Normative Cortical Thickness Trajectories along TGG2. Normative charts derived from spatial queries defined by probabilistic strata along the second growth gradient (TGG2, shown at the top left). For each of the eleven strata, the corresponding cortical delineation and age-dependent normative trajectory are displayed. Charts depict cortical thickness trajectories estimated using SNM-1000 as a function of age, after accounting for site and sex effects. Solid lines indicate the median estimate, with shaded regions showing the [25, 75], [5, 95], and [1, 99] percentiles. Scatter points represent queried cortical thickness values from held-out test subjects not used for model training. Together, these trajectories highlight the sensitivity of TGG2 to late-life thinning rates. While all strata show largely consistent characteristics in childhood, adolescence, and early adulthood, they diverge primarily after age 30: one end of the gradient shows accelerated thinning (e.g., the superior temporal and precentral gyri), whereas the other end exhibits relative stability and reduced late-life thinning (e.g., visual and ventromedial prefrontal cortices).



Extended Data Fig. E.7. Normative Cortical Thickness Trajectories along TGG3. Normative charts derived from spatial queries defined by probabilistic strata along the third growth gradient (TGG3, shown at the top left). For each of the eleven strata, the corresponding cortical delineation and age-dependent normative trajectory are displayed. Charts depict cortical thickness trajectories estimated using SNM-1000 as a function of age, after accounting for site and sex effects. Solid lines indicate the median estimate, with shaded regions showing the [25, 75], [5, 95], and [1, 99] percentiles. Scatter points represent queried cortical thickness values from held-out test subjects not used for model training. Together, these trajectories illustrate that TGG3 captures variation in early-life cortical pruning rates. Notably, the strata are differentiated by accelerated early-life thinning on one end (e.g., cingulate and ventrolateral cortex), whereas the other end exhibits slower pruning (e.g., precentral gyrus and temporal pole).



Extended Data Fig. E.8. Relationship Between Developmental Cortical Pruning and Aging-Related Thickness Decline. (A) Cortical vertex-wise centile maps for developmental cortical pruning (DCP; top), defined as the average annual rate of cortical thickness change between ages 5–20 years, and aging thickness decline (ATD; bottom), defined analogously between ages 40–95 years. **(B)** Median normative cortical thickness trajectories for eleven probabilistic strata sampled along the centiles of each map (centile targets $C = [0, 10, 20, \ldots, 100]$), estimated using SNM-1000. **(C)** Corresponding annual rates of cortical thickness percent change for each stratum. **(D)** Scatter plot showing the association between DCP and ATD across cortical vertices, revealing a significant negative relationship ($r = -0.43$, $p_{\text{spin}} = 0.017$). Together, these results indicate that cortical regions exhibiting faster pruning during development tend to show relative preservation from aging-related thinning, suggesting that spatial patterns of cortical maturation and degeneration partially mirror one another, in line with the retrogenesis hypothesis.

Supplementary material for:

Spectral normative modeling of brain structure

Sina Mansour L. [1,2,3,4,5,6] ✉, **Maria A. Di Biase** [2,7,8], **Chen Zhang** [1], **Fang Tian** [1,3,4,5,6,9,10], **Shaoshi Zhang** [1,3,4,5,6,9,10,11,12], **Hongwei Yan** [1], **Aihuiping Xue** [1], **Joanna Su Xian Chong** [1,3,4,5], **Niousha Dehestani** [1,3,4,5], **Eric Kwun-Kei Ng** [1,3,4,5], **Fang Ji** [1,3,4,5], **Xing Qian** [1,3,4,5], **Yichi Zhang** [1,3,4,5], **Wen Liang Loh** [1,3,4,5], **Joice Sin Yi Tham** [1,3,4,5], **Voon Hao Lew** [1,3,4,5], **Samantha Hui Fang Neo** [1,3,4,5], **Fiona Jia Wen Goh** [1,3,4,5], **Narayanaswamy Venketasubramanian** [13,14], **Eddie Chong** [14], **Nagaendran Kandiah** [15], **Ai Peng Tan** [16,17,18,19], **Michael J. Meaney** [20], **Marielle V. Fortier** [18], **Yap Seng Chong** [18,21], **Woon-Puay Koh** [3], **Vanessa Cropley** [22,23], **Jakob Seidlitz** [24,25,26,27], **Aaron Alexander-Bloch** [24,25,26,27], **Richard Bethlehem** [28], **Christopher Chen** [14,29], **Juan Helen Zhou** [1,3,4,5,9,11], **the Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing** *, **the Alzheimer's Disease Neuroimaging Initiative** **, **the Lifespan Brain Chart Consortium**, **B.T. Thomas Yeo** [1,3,4,5,6,9,10,11,30] ⚖, ✉, and **Andrew Zalesky** [2,31,] ⚖, ✉

[1] *Centre for Sleep & Cognition & Centre for Translational Magnetic Resonance Research, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[2] *Systems Neuroscience Lab, Department of Psychiatry, The University of Melbourne, Parkville, Victoria, Australia*
[3] *Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[4] *Human Potential Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[5] *Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[6] *Institute for Digital Medicine (WisDM), Yong Loo Lin School of Medicine, National University of Singapore, Singapore*
[7] *Stem Cell Disease Modelling Lab, Department of Anatomy and Physiology, The University of Melbourne, Parkville, Victoria, Australia*
[8] *Psychiatry Neuroimaging Laboratory, Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, USA*
[9] *Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore*
[10] *N.1 Institute for Health, National University of Singapore, Singapore, Singapore*
[11] *Integrative Sciences and Engineering Programme (ISEP), National University of Singapore, Singapore, Singapore*
[12] *Center for Brain Circuit Therapeutics, Brigham and Women's Hospital, Harvard Medical School, Boston, USA*
[13] *Raffles Neuroscience Centre, Raffles Hospital, Singapore, Singapore*
[14] *Memory Aging and Cognition Centre, National University Health System, Singapore, Singapore*
[15] *Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore*
[16] *Department of Diagnostic Imaging, National University Health System Singapore, Singapore, Singapore*
[17] *Department of Diagnostic Radiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[18] *Institute for Human Development and Potential, Agency for Science, Technology and Research in Singapore, Singapore, Singapore*
[19] *Yeo Boon Khim Mind Science Centre, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[20] *Douglas Mental Health University Institute, McGill University, Montreal, Canada*
[21] *Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[22] *Centre for Youth Mental Health, University of Melbourne, Parkville, Victoria, Australia*
[23] *Orygen Youth Health, Orygen, Parkville, Victoria, Australia*
[24] *Lifespan Brain Institute, Children's Hospital of Philadelphia, Philadelphia, PA United States*
[25] *Department of Child and Adolescent Psychiatry and Behavioral Science, Children's Hospital of Philadelphia, Philadelphia, PA United States*
[26] *Department of Psychiatry, University of Pennsylvania, Philadelphia, PA United States*
[27] *Brain-Gene Development Laboratory, Lifespan Brain Institute at Children's Hospital of Philadelphia and University of Pennsylvania, Philadelphia, PA, United States*
[28] *Department of Psychology, University of Cambridge, Cambridge, United Kingdom*
[29] *Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*
[30] *Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States*
[31] *Department of Biomedical Engineering, The University of Melbourne, Parkville, Victoria, Australia*
* *Data used in this article were obtained in part from the Australian Imaging Biomarkers and Lifestyle flagship study of ageing (AIBL). See https://aibl.org.au/ and the Acknowledgments for further details.*
** *Data used in this article were obtained in part from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). Additional details are provided in the Acknowledgments.*
⚖ *These authors contributed equally to this work.*
✉ *E-mail: sina.mansour.lakouraj@gmail.com; thomas.yeo@nus.edu.sg; azalesky@unimelb.edu.au*

Supplementary Table of Contents

S.1 Supplementary Information **III**
S.1.1 Healthy Lifespan Imaging Data . . . III
S.1.1.1 ABCD Data . . . IV
S.1.1.2 ABIDE Data . . . IV
S.1.1.3 ADHD200 Data . . . IV
S.1.1.4 ADNI Data . . . IV
S.1.1.5 AIBL Data . . . V
S.1.1.6 AOMIC Data . . . V
S.1.1.7 ASRB Data . . . V
S.1.1.8 BGSP Data . . . VI
S.1.1.9 BHRC Data . . . VI
S.1.1.10 CCNP Data . . . VI
S.1.1.11 CORR Data . . . VI
S.1.1.12 DLBS Data . . . VII
S.1.1.13 GUSTO Data . . . VII
S.1.1.14 HBN Data . . . VII
S.1.1.15 HCP Aging Data . . . VII
S.1.1.16 HCP Development Data . . . VIII
S.1.1.17 HCP Young Adult Data . . . VIII
S.1.1.18 IXI Data . . . VIII
S.1.1.19 NKI Data . . . VIII
S.1.1.20 OASIS Data . . . IX
S.1.1.21 PNC Data . . . IX
S.1.1.22 SALD Data . . . IX
S.1.1.23 SG70 Data . . . IX
S.1.1.24 SG Lifespan Data . . . X
S.1.1.25 SINGER Data . . . X
S.1.1.26 SLIM Data . . . X
S.1.1.27 SPRESTO Data . . . X
S.1.1.28 TCP Data . . . XI
S.1.1.29 UKB Data . . . XI
S.1.1.30 WAYNE Data . . . XI
S.1.2 Sample Inclusion Criteria . . . XII
S.1.3 Stratified Sample Splitting . . . XII
S.1.4 MACC Clinical Imaging Data . . . XIII
S.1.5 Cross-Basis Correlation Structure . . . XIV
S.1.6 Benchmarking High-Resolution Spatial Queries . . . XV
S.1.7 Cortical Projections of Performance Metrics . . . XV
S.1.7.1 Reconstruction Accuracy . . . XV
S.1.7.2 Reconstruction Residuals . . . XVII
S.1.7.3 Brain-Wide Normative Estimates . . . XVII
S.1.7.4 Regional Normative Estimates . . . XVIII
S.1.7.5 High-Resolution Normative Estimates . . . XIX
S.1.8 Sensitivity Analyses for Query Characteristics . . . XX
S.1.8.1 Effect of Spatial Granularity . . . XX
S.1.8.2 Effect of Spatial Asymmetry . . . XXIII
S.1.9 Within-Group Cognitive Associations . . . XXVIII
S.1.10 Brain-Wide Cognitive Associations . . . XXX
S.1.11 Association Consistency Across Diagnostic Cohorts . . . XXX
S.1.12 Heterogeneity Landscape Subgroups . . . XXXII
S.1.13 Robustness of the Heterogeneity Landscape . . . XXXII

S.1 Supplementary Information

S.1.1 Healthy Lifespan Imaging Data

Healthy human brain imaging data were aggregated from a large multi-cohort collection of publicly available neuroimaging datasets to capture normative variation across the human lifespan. In total, 30 primary datasets contributed to the healthy reference sample used for model training, encompassing a broad age range and diverse acquisition protocols, resulting in 78,405 individual scans after quality control and participant selection procedures (see Section S.1.2 Sample Inclusion Criteria). Table S.1 summarizes population demographics across the complete samples of each dataset. Two separate CSV files are provided for further details: one reporting demographic characteristics across train and test splits for each dataset (`dataset_demography.csv`), and another providing site- and scanner-specific demographic breakdowns (`cross_site_demography.csv`). The following sections provide dataset-specific details on image acquisition and preprocessing procedures for each contributing cohort.

Cohort	Sample Size	Age	Sex	Avg. Thickness	# Sites/Scanners
ABCD	9610	11.86 ± 2.23	51.1% F	2.69 ± 0.09	32
ABIDE	1003	16.01 ± 8.65	23.9% F	2.64 ± 0.17	31
ADHD200	525	12.20 ± 3.37	47.8% F	2.72 ± 0.10	10
ADNI	676	73.66 ± 7.04	54.9% F	2.39 ± 0.12	34
AIBL	472	73.01 ± 6.20	58.9% F	2.42 ± 0.10	3
AOMIC	1357	22.60 ± 1.77	53.9% F	2.62 ± 0.09	3
ASRB	185	40.54 ± 14.03	50.8% F	2.58 ± 0.09	5
BGSP	1557	21.51 ± 2.86	57.6% F	2.49 ± 0.07	4
BHRC	580	9.91 ± 1.82	41.7% F	2.67 ± 0.10	2
CCNP	194	11.92 ± 3.15	52.6% F	2.55 ± 0.10	1
CORR	1436	25.57 ± 15.02	51.3% F	2.54 ± 0.13	29
DLBS	307	53.90 ± 19.83	62.2% F	2.49 ± 0.14	1
GUSTO	535	7.40 ± 2.24	50.8% F	2.82 ± 0.09	1
HBN	1337	11.36 ± 3.75	37.5% F	2.89 ± 0.18	4
HCP-A	710	59.56 ± 14.89	55.6% F	2.58 ± 0.15	4
HCP-D	652	14.44 ± 4.06	53.8% F	3.01 ± 0.14	4
HCP-YA	1113	28.80 ± 3.70	54.4% F	2.66 ± 0.08	1
IXI	560	48.74 ± 16.48	55.7% F	2.56 ± 0.16	1
NKI	1287	39.88 ± 21.73	60.7% F	2.64 ± 0.17	1
OASIS	878	70.32 ± 8.49	58.8% F	2.41 ± 0.09	3
PNC	1563	15.04 ± 3.66	52.8% F	2.69 ± 0.10	1
SALD	475	45.07 ± 17.44	62.9% F	2.55 ± 0.10	1
SG70	978	74.83 ± 2.60	54.3% F	2.31 ± 0.12	1
SG Lifespan	240	39.98 ± 14.91	61.2% F	2.48 ± 0.10	1
SINGER	920	68.41 ± 4.80	51.8% F	2.34 ± 0.10	2
SLIM	570	20.16 ± 1.35	56.0% F	2.61 ± 0.08	1
SPRESTO	127	5.24 ± 0.52	53.5% F	2.77 ± 0.08	1
TCP	91	32.83 ± 13.13	54.9% F	2.39 ± 0.10	2
UKB	48272	65.63 ± 7.62	49.8% F	2.69 ± 0.11	4
WAYNE	195	51.39 ± 15.42	66.7% F	2.49 ± 0.08	1

Table S.1. Sample Characteristics of the Healthy Reference Dataset. Values are presented as $\mu \pm \sigma$, indicating the mean (μ) and standard deviation (σ). The Sites / Scanners column indicates the number of unique site/scanner combinations per cohort.

S.1.1.1 ABCD Data

The Adolescent Brain Cognitive Development® (ABCD) Study is a landmark, longitudinal investigation of brain development and child health, collected across 21 sites throughout the United States [1]. Structural MRI data were acquired on three 3T scanner platforms (Siemens Prisma, General Electric 750, and Philips) and T1-weighted images were preprocessed using an automated FreeSurfer pipeline (v7.1.1). Preprocessed FreeSurfer outputs were accessed through the NIH Brain Development Cohorts Data Sharing Platform (NBDC). For inclusion in the healthy reference sample, participants were required to have no history of neurological, psychiatric, developmental, or major medical conditions. Specifically, participants were excluded if parental responses indicated a history of neurological or medical disorders (e.g., cerebral palsy, brain tumor, stroke, epilepsy), psychiatric or developmental conditions (e.g., ADHD, autism spectrum disorder, schizophrenia, intellectual disability, mood or anxiety disorders), or substance use disorders, based on the ABCD screening and risk questionnaire (`https://nda.nih.gov/data_structure.html?short_name=abcd_screen01`). Additional quality control procedures were applied as detailed in Section S.1.2 Sample Inclusion Criteria. After applying these exclusion criteria, the final reference sample comprised 9,610 individuals (51.1% female) aged 8 to 18 years, spanning 32 unique site/scanner combinations.

S.1.1.2 ABIDE Data

The Autism Brain Imaging Data Exchange (ABIDE) initiative has aggregated functional and structural neuroimaging data collected from laboratories worldwide to accelerate research into the neural bases of autism spectrum disorder (ASD)[2,3]. In this study, T1-weighted structural MRI data from ABIDE were used. All structural images were processed using an automated pipeline running `recon-all` as implemented in FreeSurfer (v5.3.0). To construct a healthy reference sample, individuals with a reported ASD diagnosis were excluded. After applying additional quality control and participant selection procedures (see Section S.1.2 Sample Inclusion Criteria), the final ABIDE sample comprised 1,003 individuals (23.9% female) spanning ages 6 to 64 years. These data were acquired across 31 distinct site/scanner combinations, reflecting the multi-site and multi-scanner nature of the ABIDE data.

S.1.1.3 ADHD200 Data

The ADHD200 Sample is a grassroots, open data-sharing initiative designed to accelerate understanding of the neural basis of attention-deficit/hyperactivity disorder (ADHD) through large-scale, multi-site neuroimaging aggregation and discovery-based science. The initiative publicly released resting-state functional MRI and T1-weighted structural MRI data collected across multiple independent imaging sites, comprising both typically developing individuals and children and adolescents diagnosed with ADHD [4]. For the present study, only T1-weighted structural MRI data from typically developing participants were considered for inclusion in the healthy reference sample. Structural images were processed using FreeSurfer (v6.0.0) to derive cortical thickness measures, following additional quality control and participant selection procedures (see Section S.1.2 Sample Inclusion Criteria). After exclusion of individuals with an ADHD diagnosis and scans failing quality control, the final ADHD200 sample comprised 525 individuals (47.8% female), spanning ages 7 to 25 years, and distributed across 10 distinct site/scanner combinations.

S.1.1.4 ADNI Data

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database [5,6]. ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The original goal of ADNI was to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). The current goals include validating biomarkers for clinical trials, improving the generalizability of ADNI data by increasing diversity in the participant cohort, and to provide data concerning the diagnosis and progression of Alzheimer's

disease to the scientific community. For up-to-date information, see https://adni.loni.usc.edu. To date, ADNI has comprised three major phases, ADNI1 (2004–2009), ADNI-GO/2 (2010–2016), and ADNI3 (2017–2023), each incorporating newly enrolled participants as well as individuals continuing from earlier phases. Across these phases, MRI data were acquired using a range of scanner models and imaging protocols, with ADNI1 primarily employing 1.5T scanners and later phases predominantly adopting 3T scanners, contributing to substantial site- and scanner-level heterogeneity.

For the present study, only T1-weighted structural MRI data from cognitively normal (CN) participants were included in the healthy reference sample. Individuals diagnosed with mild cognitive impairment (MCI) or Alzheimer's disease were excluded. T1-weighted images were obtained from the USC Laboratory of Neuro Imaging (LONI) Image and Data Archive, with accompanying phenotypic and diagnostic information extracted from the ADNIMERGE dataset[7]. All structural images were processed using FreeSurfer (v6.0.0) to derive cortical thickness measures, following standard automated reconstruction procedures and additional quality control steps (see Section S.1.2 Sample Inclusion Criteria). After applying all inclusion, exclusion, and quality control criteria, the final ADNI sample comprised 676 cognitively normal individuals (54.9% female), spanning ages 50 to 93 years, and distributed across 34 distinct site/scanner combinations.

S.1.1.5 AIBL Data

The Australian Imaging, Biomarkers and Lifestyle Flagship Study of Ageing (AIBL) is a longitudinal research initiative designed to identify biomarkers, cognitive characteristics, and lifestyle factors associated with healthy aging and the development of Alzheimer's disease [8]. Data were collected by the AIBL study group, and the study methodology has been described previously [9]. For the present study, only T1-weighted structural MRI data from cognitively normal participants (CN) were included. Participants with mild cognitive impairment, Alzheimer's disease, or missing/uncertain diagnostic status (codes -4, 7, or NaN) were excluded. Structural images were processed using FreeSurfer (v6.0.0) to derive cortical thickness measures, following additional quality control and participant selection procedures (see Section S.1.2 Sample Inclusion Criteria). The final AIBL sample comprised 474 individuals (58.9% female), aged 60 to 92 years, across 3 distinct site/scanner combinations.

S.1.1.6 AOMIC Data

The Amsterdam Open MRI Collection (AOMIC) is a multi-cohort initiative providing multimodal MRI data, including structural T1-weighted, diffusion-weighted, and functional (resting-state and task-based) MRI, alongside detailed demographics and psychometric assessments [10]. AOMIC comprises three datasets: ID1000, PIOP1, and PIOP2. All data were acquired on the same Philips 3T scanner (Philips, Best, The Netherlands), with scanner hardware and software upgrades occurring between studies. Scans were performed using a 32-channel head coil, which was upgraded alongside the dStream version of the scanner for PIOP2. Structural T1-weighted images were processed using FreeSurfer (v6.0.0) to derive cortical thickness measures. For inclusion in the healthy reference sample, only participants with complete data and passing quality control were retained (see Section S.1.2 Sample Inclusion Criteria). The final AOMIC sample consisted of 1,357 healthy individuals (53.9% female), aged 18 to 26 years.

S.1.1.7 ASRB Data

The Australian Schizophrenia Research Bank (ASRB) is a multi-site biobank providing clinical, neuroimaging, and genetic data from individuals with schizophrenia and healthy comparison participants [11]. Structural T1-weighted MRI scans were acquired using Siemens Avanto scanners across five sites in Australia (Melbourne, Sydney, Brisbane, Perth, and Newcastle), employing a harmonized MPRAGE protocol (1 mm isotropic voxels, TR = 1980 ms, TE = 4.3 ms). To quantify inter-site variability, a single individual was scanned at all sites, and a Siemens MRI phantom was used for calibration prior to image acquisition. For the present study, only healthy participants without a diagnosis were included, following additional quality control procedures to exclude scans with gross artifacts or significant motion (see Section S.1.2 Sample Inclusion Criteria). T1-weighted images were processed using

FreeSurfer (v6.0.0) without manual edits of cortical surfaces. The final ASRB sample comprised 185 individuals (50.8% female), aged 18 to 65 years, distributed across five site/scanner combinations.

S.1.1.8 BGSP Data

The Brain Genomics Superstruct Project (BGSP) Open Access Data Release[12,13] comprises a carefully-vetted collection of neuroimaging, behavioral, cognitive, and personality data from over 1,500 healthy participants. Each neuroimaging dataset includes one T1-weighted structural MRI acquisition and one or more resting-state functional MRI acquisitions, with accompanying automated quality assessments and pre-computed brain morphometrics. For the present study, T1-weighted structural scans were processed using FreeSurfer (v6.0.0) to derive cortical thickness measures. After additional quality control and participant selection procedures (see Section S.1.2 Sample Inclusion Criteria), the final BGSP sample comprised 1,557 individuals (57.6% female), aged 19 to 35 years, spanning 4 distinct site/scanner combinations.

S.1.1.9 BHRC Data

The Brazilian High-Risk Cohort Study for Childhood Psychiatric Disorders (BHRC) is a large, community-based longitudinal study designed to investigate typical and atypical developmental trajectories of psychopathology and cognition in childhood and adolescence[14]. Participants were recruited from state-funded schools in the Brazilian cities of São Paulo and Porto Alegre, comprising both a representative community sample and children at increased familial risk for mental disorders. Neuroimaging data were acquired across two sites using consistent acquisition protocols, with participants re-scanned on the same scanner across follow-up waves. For the present study, BHRC data were accessed indirectly through the Reproducible Brain Charts (RBC) initiative, which harmonizes and redistributes curated neuroimaging datasets to facilitate large-scale normative modeling[15]. RBC-provided structural MRI derivatives, preprocessed using FreeSurfer (v6.0.1) and sMRIPrep[16] (v0.7.1), were retrieved using `datalad`[17]. Following quality control and exclusion procedures (see Section S.1.2 Sample Inclusion Criteria), the final BHRC sample comprised 580 healthy individuals (41.7% female), aged 5 to 15 years, spanning 2 distinct site/scanner combinations.

S.1.1.10 CCNP Data

The Chinese Color Nest Project (CCNP) is a large-scale, longitudinal neuroimaging initiative with the long-term goal of characterizing normative developmental trajectories of brain structure and function across the lifespan, and of understanding neurobiological changes associated with the emergence of mental health and learning disorders [18,19]. CCNP comprises three coordinated phases spanning development (devCCNP), maturation (matCCNP), and aging (ageCCNP). The present study makes use of data from the developmental CCNP (devCCNP) cohort, also referred to as "Growing Up in China", which includes longitudinal neuroimaging data from typically developing children and adolescents aged 6 to 18 years. For the present study, CCNP data were accessed indirectly through RBC [15]. RBC-provided structural MRI derivatives, preprocessed using FreeSurfer (v6.0.1) and `sMRIPrep`[16] (v0.7.1), were retrieved using `datalad`[17]. Following quality control and exclusion procedures (see Section S.1.2 Sample Inclusion Criteria), the final CCNP sample comprised 194 healthy individuals (52.6% female), aged 6 to 18 years, collected at a single imaging site.

S.1.1.11 CORR Data

CORR Data were obtained from the Consortium for Reliability and Reproducibility (CoRR), an open neuroimaging resource established to facilitate the assessment of test–retest reliability and reproducibility in functional and structural connectomics[20]. The present study included 1,436 individuals, aged 6–88 years (51.3% female), acquired across 29 site/scanner combinations. As CoRR aggregates retrospectively collected datasets from multiple international laboratories, the sample spans a wide range of acquisition protocols and study designs. All structural T1-weighted MRI scans were processed using an automated pipeline running `recon-all` as implemented in FreeSurfer (v5.3.0).

S.1.1.12 DLBS Data

The Dallas Lifespan Brain Study (DLBS) is a longitudinal, open neuroimaging cohort designed to characterize cognitive and brain changes across the adult lifespan, with a particular emphasis on trajectories of healthy aging [21,22]. The study includes extensively phenotyped participants spanning early to late adulthood, with comprehensive assessments of cognition, brain structure, and brain function. Neuroimaging data were acquired at a single imaging site using a 3T MRI scanner, ensuring consistency of acquisition across participants and study waves. For the present study, structural T1-weighted MRI data were used. Raw structural images were processed using FreeSurfer (v6.0.0) to derive cortical morphometric measures. Following exclusion of individuals failing quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final DLBS sample comprised 307 healthy individuals (62.2% female), aged 20 to 89 years.

S.1.1.13 GUSTO Data

The Growing Up in Singapore Towards Healthy Outcomes (GUSTO) study is a prospective, longitudinal birth cohort established to investigate early-life determinants of health and neurodevelopment in children growing up in Singapore [23]. The cohort recruited pregnant women attending first-trimester antenatal ultrasound clinics at the National University Hospital and KK Women's and Children's Hospital. Parents were Singapore citizens or permanent residents of Chinese, Malay, or Indian ethnic background. Birth outcomes and pregnancy-related measures were obtained from hospital records, and socioeconomic indicators, including household income, were collected via structured questionnaires administered during pregnancy. The GUSTO study was approved by the National Healthcare Group Domain Specific Review Board and the SingHealth Centralized Institutional Review Board, and written informed consent was obtained from all participating mothers. Children underwent MRI scanning at multiple time points across development. For the present study, structural T1-weighted MRI scans were processed using FreeSurfer (v6.0.0) to derive cortical morphometric measures. Following quality control and exclusion procedures (see Section S.1.2 Sample Inclusion Criteria), the final GUSTO sample comprised 535 healthy children (50.8% female), aged 4 to 11 years.

S.1.1.14 HBN Data

The Healthy Brain Network (HBN) is a large-scale, ongoing initiative led by the Child Mind Institute, designed to create and openly share a richly phenotyped biobank to advance understanding of brain development and developmental psychopathology in childhood and adolescence [24]. The HBN cohort comprises children and adolescents recruited from the New York City area, including both typically developing and help-seeking individuals, and includes extensive psychiatric, behavioral, cognitive, and lifestyle assessments alongside multimodal neuroimaging data. Neuroimaging protocols include structural T1-weighted MRI, diffusion MRI, and functional MRI acquired across multiple imaging sites. For the present study, HBN data were accessed indirectly through RBC [15]. RBC-provided structural MRI derivatives, preprocessed using FreeSurfer (v6.0.1) and `sMRIPrep`[16] (v0.7.1), were retrieved using `datalad`[17]. Following additional exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final HBN sample comprised 1,337 healthy individuals (37.5% female), aged 5 to 22 years, collected across four imaging sites.

S.1.1.15 HCP Aging Data

The HCP Aging (HCP-A) imaging data was sourced from the HCP-Aging Lifespan 2.0 release[25] (HCP-A) which included structural MRI images from 725 healthy adults aged 36 to 100+ years old. For all participants aged above 90 years old, age was reported as 100 to avoid releasing protected health information. Participants with ages above 90 were excluded from the analyses due to the lack of precise age information, resulting in a final sample of 710 individuals (55.6% female). The acquisition and preprocessing protocols for this cohort were generally consistent with the earlier acquired HCP-YA protocols (see Section S.1.1.17 HCP Young Adult Data), with particular necessary changes due to hardware differences that are detailed elsewhere[26]. In summary, data were collected on Siemens 3T Prisma scanners (with 32-channel head coils) located at 4 different acquisition sites in the United

States. Particularly, T1w images were collected by a multi-echo MPRAGE sequence with real-time motion correction (0.8mm isotropic voxels, TR = 2500ms, TE = 1.8/3.6/5.4/7.2ms, TI = 1000ms, flip angle = 8°, 6/8 slice partial Fourier)[26]. The preprocessing pipeline for structural images was similar to that of HCP-YA following the minimal preprocessing pipeline[27] with minor variations including a different version of FreeSurfer (version 6.0) and an update to MSMAll using higher order smoothness constraints[28].

S.1.1.16 HCP Development Data

The development imaging data was sourced from HCP-Development Lifespan 2.0 release[29] (HCP-D) which contained structural MRI images of 652 healthy participants (53.8% female) aged 5 to 22 years old. This cohort aimed to adapt existing HCP protocols to the practical challenges of studying developmental populations. The acquisition and processing of the HCP-D dataset[29] were generally consistent with the HCP-A dataset[25]. Namely, similar Prisma scanners (with 32-channel head coils) were used to collect imaging data across four acquisition sites, while a tailored 32-channel head coil was used for 5-7 year-old participants[26,29]. Identical to HCP-A, a multi-echo MPRAGE sequence was used to collect T1w images followed by the same preprocessing procedures[26].

S.1.1.17 HCP Young Adult Data

Young adult imaging data was sourced from the HCP S1200 release[30] (HCP-YA). This contained structural MRI images obtained from 1096 healthy participants (54.4% female) aged 22 to 37 years old. A comprehensive report on imaging acquisition and preprocessing is available elsewhere[27]. In brief, the images were acquired with a Siemens 3T Skyra scanner with a 32-channel head coil (housed at Washington University). T1w images were acquired using a single-echo MPRAGE sequence (0.7mm isotropic resolution, TR = 2400ms, TE = 2.14ms, TI = 1000ms, flip angle = 8°, no partial Fourier)[27]. Acquired images underwent preprocessing using the HCP's minimal preprocessing pipeline. Notably, structural images underwent gradient distortion correction, skull stripping, readout distortion correction, and bias field correction, followed by a nonlinear registration to a standard template volumetric space. Cortical surface reconstruction was performed using FreeSurfer (version 5.2) which produced 3-dimensional meshes for the white and pial cortical surface along with segmentation and folding-based registration to the *fsaverage* surface. Finally, the outputs were registered to the *Conte69* surface template with a multimodal surface matching algorithm (MSMAll)[31] and downsampled to triangulated meshes with 32k vertices per cortical hemisphere, i.e. the *fs-LR* 32k space.

S.1.1.18 IXI Data

The IXI dataset is an open-access neuroimaging resource comprising structural and multimodal MRI data from nearly 600 healthy adult participants, made publicly available as part of the Information eXtraction from Images (IXI) project (`https://brain-development.org/ixi-dataset/`). MRI data were collected at three hospitals in London using different scanners, including a Philips 3T system at Hammersmith Hospital, a Philips 1.5T system at Guy's Hospital, and a GE 1.5T system at the Institute of Psychiatry. The dataset includes T1-, T2-, and proton density–weighted images, as well as diffusion and angiographic sequences. For the present study, only T1-weighted structural MRI scans were considered for inclusion. Structural images were processed using FreeSurfer (v6.0.1) to derive cortical morphometric measures. Following quality control and participant selection procedures (see Section S.1.2 Sample Inclusion Criteria), the final IXI sample comprised 560 healthy individuals (55.7% female), aged 20 to 86 years.

S.1.1.19 NKI Data

The Nathan Kline Institute (NKI) Rockland Sample is a large-scale, community-ascertained neuroimaging initiative designed to characterize normative and atypical brain organization across the human lifespan [32]. The study aims to provide a deeply phenotyped sample spanning childhood through late adulthood, including detailed demographic, behavioral, and neuroimaging measures. Participants

were recruited from the United States and assessed under protocols approved by the Institutional Review Board at the Nathan Kline Institute. For the present study, NKI data were accessed indirectly through RBC [15]. RBC-provided structural MRI derivatives, preprocessed using FreeSurfer (v6.0.1) and `sMRIPrep`[16] (v0.7.1), were retrieved using `datalad`[17]. Following additional exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final NKI sample comprised 1,287 healthy individuals (60.7% female), aged 6 to 86 years, collected at a single imaging site.

S.1.1.20 OASIS Data

The Open Access Series of Imaging Studies (OASIS) is a multimodal neuroimaging resource designed to facilitate the study of healthy aging and Alzheimer's disease across the adult lifespan [33]. The OASIS initiative aggregates data from several longitudinal and cross-sectional studies conducted through the Charles F. and Joanne Knight Alzheimer Disease Research Center (Knight ADRC) at Washington University in St. Louis, including the Memory and Aging Project, the Adult Children Study, and the Healthy Aging and Senile Dementia study. Imaging data were acquired over multiple years using both 1.5T and 3T MRI scanners across several scanner platforms and acquisition protocols. The present study makes use of data from the OASIS-3 release, which integrates earlier OASIS cohorts into a large, longitudinal dataset spanning mid- to late adulthood. For inclusion in the healthy reference sample, only cognitively normal individuals were retained, excluding participants with Alzheimer's disease or other forms of dementia. MRI scans and accompanying phenotypic data were downloaded from XNAT Central [34]. Structural T1-weighted MRI scans were processed using FreeSurfer (v6.0.0) to derive cortical thickness measures, following additional quality control and participant selection procedures (see Section S.1.2 Sample Inclusion Criteria). The final OASIS sample comprised 878 healthy individuals (58.8% female), aged 42 to 96 years, distributed across 3 distinct site/scanner combinations.

S.1.1.21 PNC Data

The Philadelphia Neurodevelopmental Cohort (PNC) is a large, community-based neuroimaging initiative designed to characterize brain development and its relationship to cognition, behavior, and mental health across adolescence and young adulthood [35,36]. Participants were recruited from the Children's Hospital of Philadelphia care network, forming a diverse developmental sample of children and adolescents residing in the greater Philadelphia area, United States. The study was approved by the Institutional Review Boards of the University of Pennsylvania and the Children's Hospital of Philadelphia. For the present study, PNC data were accessed indirectly through RBC [15]. RBC-provided structural MRI derivatives, preprocessed using FreeSurfer (v6.0.1) and `sMRIPrep`[16] (v0.7.1), were retrieved using `datalad`[17]. Following additional exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final PNC sample comprised 1,563 healthy individuals (52.8% female), aged 8 to 23 years, collected at a single imaging site.

S.1.1.22 SALD Data

The Southwest University Adult Lifespan Dataset (SALD) is a cross-sectional neuroimaging cohort designed to characterize structural and functional brain changes across adulthood[37]. The dataset comprises a community-based adult lifespan sample and includes multimodal neuroimaging and basic phenotypic information. Data were collected at a single imaging site at the Center for Brain Imaging, Southwest University, under locally approved research protocols. For the present study, structural T1-weighted MRI data from SALD were used. All scans were processed using FreeSurfer (v6.0.0). Following additional exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final SALD sample comprised 475 healthy individuals (62.9% female), aged 19 to 80 years, collected at a single site.

S.1.1.23 SG70 Data

The SG70 study was conducted as a subcohort study nested within the Singapore Chinese Health Study [38] to facilitate an in-depth investigation of multi-organ and multi-system aging among a sub-

group of older adults in their seventies. In total, 1,140 participants aged 68 to 82 years were enrolled during follow-up assessments from 2022 to 2025 [39]. Participants underwent comprehensive MRI brain imaging and cognitive assessments covering multiple domains at a single study location. For the current analysis, T1-weighted structural MRI scans were processed using FreeSurfer (v5.3.0). After implementing exclusion and quality control protocols (see Section S.1.2 Sample Inclusion Criteria), data from 978 individuals (54.3% female) from the SG70 study were included in this analysis. The study received approval from both the National University of Singapore Institutional Review Board (NUS-IRB, 2020/398) and the Domain Specific Review Board (DSRB, 2020/01465), and was conducted in accordance with the Human Biomedical Research Act (HBRA) and ICH-GCP guidelines.

S.1.1.24 SG Lifespan Data

The SG Lifespan dataset is a multimodal, community- and staff-based cohort established by the Multimodal Neuroimaging in Neuropsychiatric Disorders Laboratory (MNNDL) to investigate how brain structure, activity, and connectivity underlie perception and cognition across the lifespan. Data collection includes neuroimaging, physiological and psychological characteristics, cognitive and behavioral assessments, blood biomarkers, and digital health tracking (e.g., sleep and activity monitoring) across different age cohorts. The study recruited three sub-cohorts: young adults (21–30 years) and middle-aged adults (40–80 years) from the community, and NUS staff participants referred via the NUS1000 Staff Edition study in collaboration with the Sleep and Cognition Lab. For the present study, structural T1-weighted MRI scans were processed using FreeSurfer (v5.3.0). Following exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final SG Lifespan sample comprised 240 healthy individuals (61.2% female), aged 21 to 73 years, all scanned at a single imaging site.

S.1.1.25 SINGER Data

The SINGER (Singapore Intervention Study to Prevent Cognitive Impairment and Disability) study is a multi-center, two-year clinical trial evaluating the effects of multidomain lifestyle interventions on cognitive decline and brain health in older adults at risk of dementia [40]. It represents Singapore's adaptation of the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) trial, following the same multidomain dementia-prevention framework but tailored to an Asian population. For the present study, structural T1-weighted MRI scans were processed using FreeSurfer (v5.3.0). After exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final sample comprised 920 healthy older adults (51.8% female), aged 60–80 years, scanned across two imaging sites: Nanyang Technological University and the National University of Singapore. The study was approved by the National Healthcare Group Domain-Specific Review Board and conducted in accordance with Good Clinical Practice guidelines, and is registered on ClinicalTrials.gov (NCT05007353).

S.1.1.26 SLIM Data

The Southwest University Longitudinal Imaging Multimodal (SLIM) dataset is a community-based cohort designed to map brain structure and function and their associations with behavior in young adults [41]. The dataset includes multimodal MRI scans, structural, diffusion, and resting-state functional MRI, alongside detailed behavioral assessments covering demographic, cognitive, and emotional measures, collected over a long-term retest duration of approximately three and a half years. Structural T1-weighted scans were processed using FreeSurfer (v5.3.0). Following exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final SLIM sample comprised 570 healthy individuals (56.0% female), aged 16 to 28 years, collected at a single imaging site.

S.1.1.27 SPRESTO Data

The Singapore PREconception Study of long-Term maternal and child Outcomes (SPRESTO) is a large, prospective, longitudinal, multi-wave observational birth cohort designed to investigate the influence of preconception and early-life factors on maternal and child health outcomes [42,43]. Women

planning pregnancy were recruited in Singapore between 2015 and 2017 through KK Women's and Children's Hospital and community outreach, and mother–child dyads have been followed longitudinally from preconception through early childhood with extensive phenotypic, clinical, and developmental assessments. Neuroimaging data were acquired as part of longitudinal follow-up assessments during early childhood. For the present study, we made use of structural MRI data collected at the preschool-age follow-up. Structural T1-weighted scans were processed using FreeSurfer. Following exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final SPRESTO sample comprised 127 healthy children (53.5% female), aged 4 to 7 years, collected at a single imaging site.

S.1.1.28 TCP Data

The Transdiagnostic Connectome Project (TCP) is a richly phenotyped neuroimaging dataset designed to investigate brain organization and transdiagnostic dimensions of psychopathology across adulthood [44]. The study includes comprehensive behavioral, cognitive, and clinical assessments alongside multimodal MRI data and was conducted across two imaging sites in the United States. MRI data were acquired using harmonized Siemens 3T Prisma scanners with standardized acquisition protocols across sites. For the present study, we made use of structural T1-weighted MRI data from the healthy comparison group. Structural scans were processed using FreeSurfer. Following exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final TCP sample comprised 91 healthy individuals (54.9% female), aged 18 to 68 years, collected across two imaging sites (Yale University and McLean Hospital).

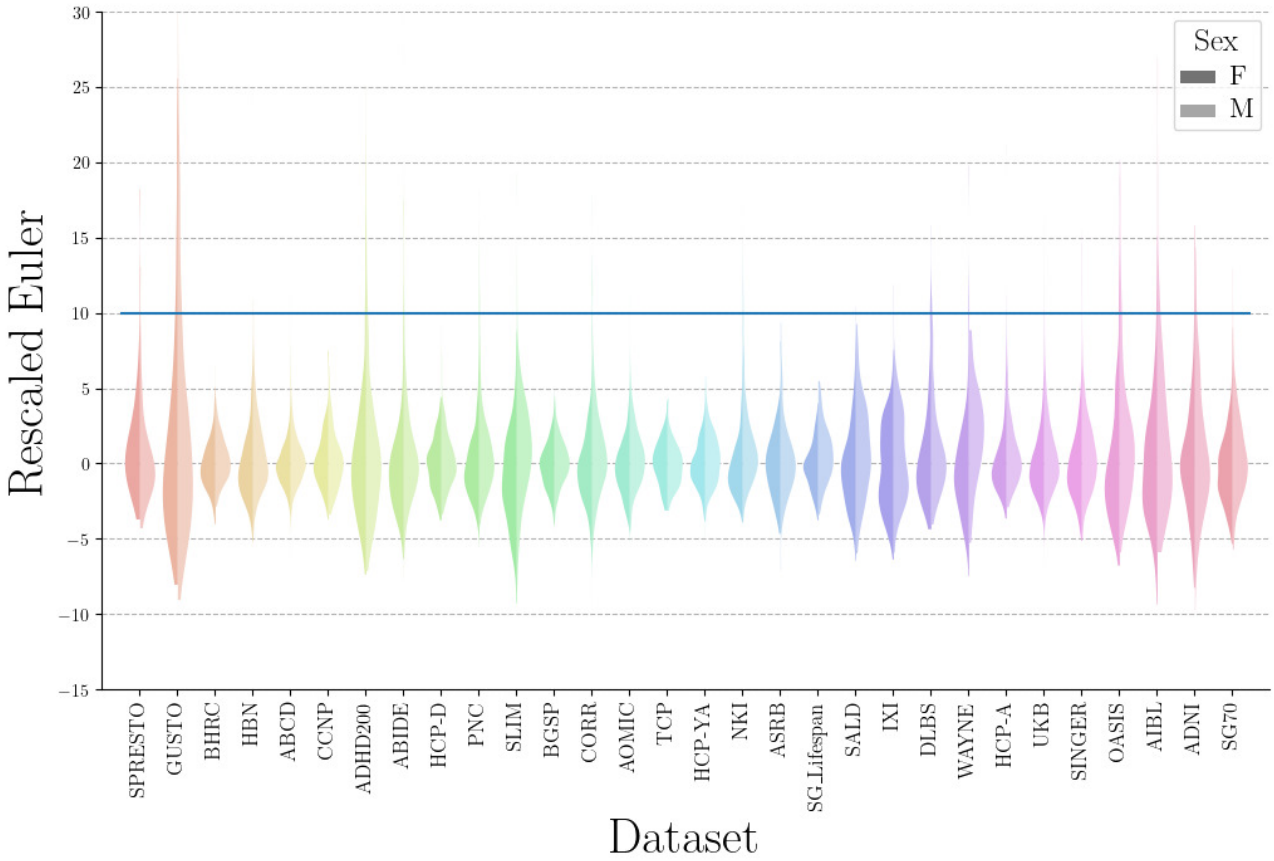
S.1.1.29 UKB Data

The UK Biobank (UKB) is a large-scale population-based cohort providing extensive phenotypic, genetic, and neuroimaging data from middle-aged and older adults in the United Kingdom [45,46]. Neuroimaging data were acquired across four scanner–site combinations using harmonized protocols [47]. For the present study, we used structural MRI data accessed from UKB application `60698`. Specifically, we made use of preprocessed FreeSurfer-derived structural outputs provided directly by UK Biobank showcase (compressed bulk file, Data-Field 20263), which include T1-weighted cortical surface models and associated structural segmentations generated using standardized UKB processing pipelines. Individuals were included as healthy participants based on responses to the UKB mental health questionnaire (Data-Field 20544), retaining only those who reported no history of mental health disorders diagnosed by a health professional. After exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final UKB sample comprised 48,272 individuals (49.8% female), aged 44 to 86 years, collected across four scanner–site combinations.

S.1.1.30 WAYNE Data

The Wayne State dataset was obtained from the Brain Aging in Detroit Longitudinal Study and made available through the 1000 Functional Connectomes Project (`https://fcon_1000.projects.nitrc.org/indi/retro/wayne_11.html`). This longitudinal study was designed to investigate brain and cognitive changes across the adult lifespan, with a focus on identifying factors influencing age-related trajectories of brain structure and function[48]. Participants were screened at baseline using a health questionnaire to exclude individuals with cardiovascular, neurological, psychiatric, endocrine, or thyroid disorders, use of centrally acting medications, excessive alcohol consumption, head injury with loss of consciousness, and other factors that could confound brain aging effects. Cognitive screening included the Center for Epidemiologic Studies Depression Scale (CES-D) and Mini-Mental State Examination (MMSE) to exclude participants with significant depressive symptomatology or cognitive impairment. Neuroimaging data were acquired at a single site using a 1.5T Siemens Magnetom scanner, with T1-weighted MPRAGE anatomical images collected across multiple waves of data acquisition. For the present study, structural T1-weighted MRI scans were processed using FreeSurfer (v6.0.0). Following exclusion and quality control procedures (see Section S.1.2 Sample Inclusion Criteria), the final WAYNE sample comprised 195 healthy individuals (66.7% female), aged 18 to 78 years, collected at a single imaging site.

S.1.2 Sample Inclusion Criteria

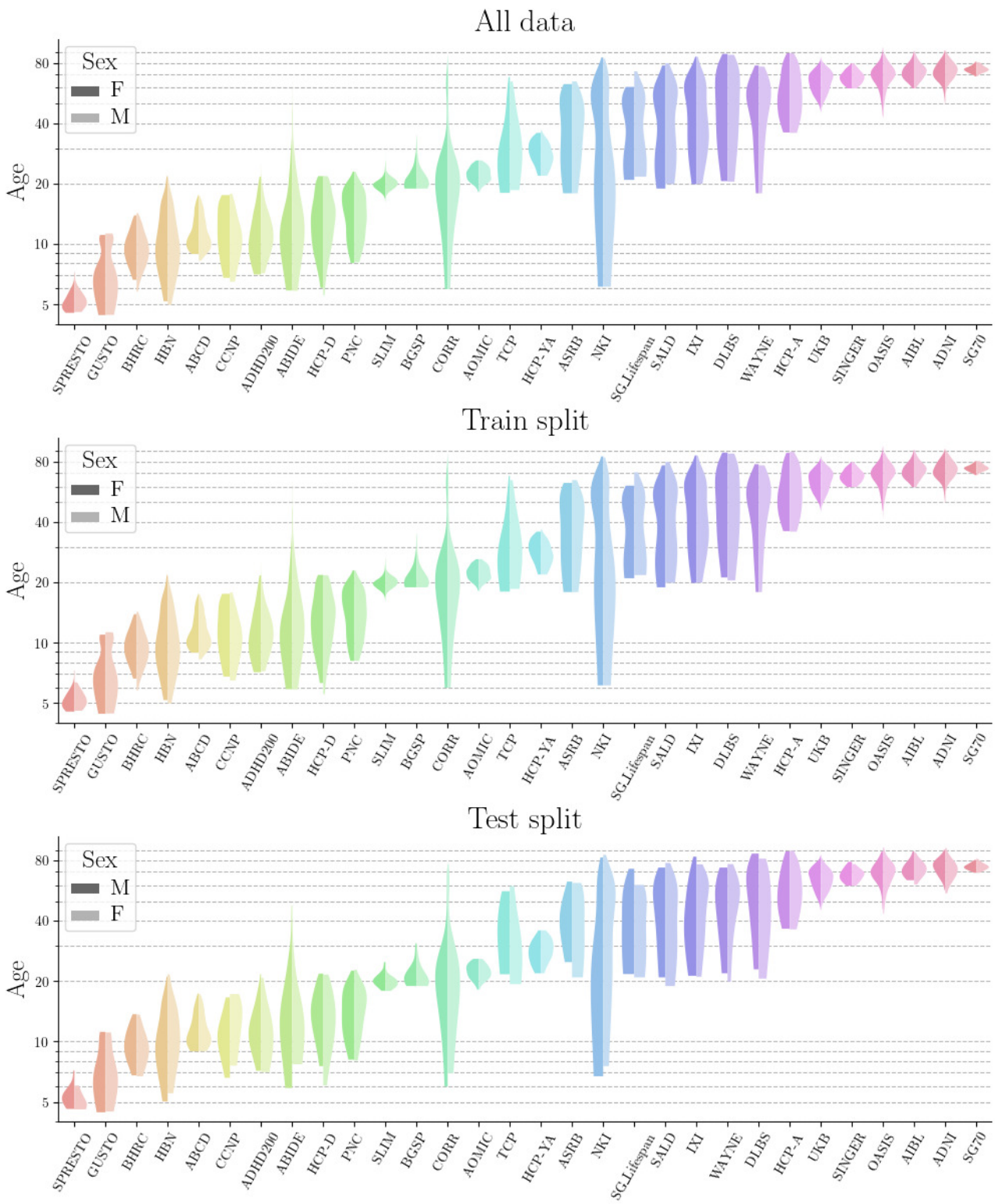


Supplementary Fig. S.1. Quality Control of Cortical Surface Reconstruction. Violin plots show the distribution of rescaled Euler numbers across datasets. For each scan, the negative Euler number was square-root transformed and median-centered within dataset. Higher values indicate lower surface reconstruction quality. Each violin corresponds to one dataset. An exclusion threshold of 10 (dashed horizontal line) was applied to remove low-quality reconstructions, following established recommendations.

Following dataset-specific exclusions (e.g., removal of individuals with clinical diagnoses where applicable), additional quality control and exclusion steps were applied. For datasets with multiple scans per individual (i.e., longitudinal follow-ups, such as ABCD, GUSTO, and UKB), a single scan per subject was randomly selected to maintain an inclusive age range while preventing overrepresentation of participants with repeated measures. Scanner–site combinations comprising fewer than 15 participants were excluded to ensure reliable estimation of site-specific effects. In line with previously established recommendations [49,50], low-quality surface reconstructions were further filtered based on the Euler number: for each scan, the negative Euler number was square-root–transformed and median-centered within each site, and scans with a resulting rescaled Euler number greater than 10 were excluded (see Supplementary Fig. S.1). This conservative, automated quality control procedure resulted in a final healthy sample of $N_p = 78,405$ individuals, which was subsequently divided into training and test splits.

S.1.3 Stratified Sample Splitting

As described in the Methods section, train test splitting was performed by a randomized approach that controls for the distribution of sample covariates (age, sec, and dataset/batch). In particular, this was implemented by `scikit-learn`'s `train_test_split` function by utilizing the `stratify` parameter. First, continuous values for individual age were grouped into bins by Quantile-based discretization. Next, discrete values encoding age, sex, and dataset information were utilized for stratified splitting. Supplementary Fig. S.2 demonstrates the stability of covariate distribution after train test splitting.



Supplementary Fig. S.2. Distribution of Demographic Information Before and After Splitting. The distribution of demographic information is displayed for the entire sample before splitting and after splitting into training and test sets. The violin plots illustrate the age distribution (y-axis) across different datasets (x-axis). Each plot differentiates between female (dark shades) and male (bright shades) participants. The stratified group splitting ensures that the demographic distributions of covariates are matched between the training and test samples.

S.1.4 MACC Clinical Imaging Data

The clinical data (part of MACC harmonization study*) was acquired by the Memory, Ageing & Cognition Centre at the National University of Singapore[51,52]. The dataset includes 542 participants aged 50 to 91 years (61.4% female, see Table S.2 for detail), grouped into three cohorts: i) a healthy

*Access to the MACC clinical imaging dataset is subject to a data-transfer agreement with the Memory, Ageing & Cognition Centre, National University of Singapore (http://www.macc.sg/).

cohort (HC) of 132 cognitively healthy individuals, ii) a mild cognitive impairment cohort (MCI) of 202 individuals with cognitive impairment but no AD, and iii) an AD cohort of 208 individuals with an AD diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria. AD diagnosis followed internationally accepted criteria using the National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) guidelines.

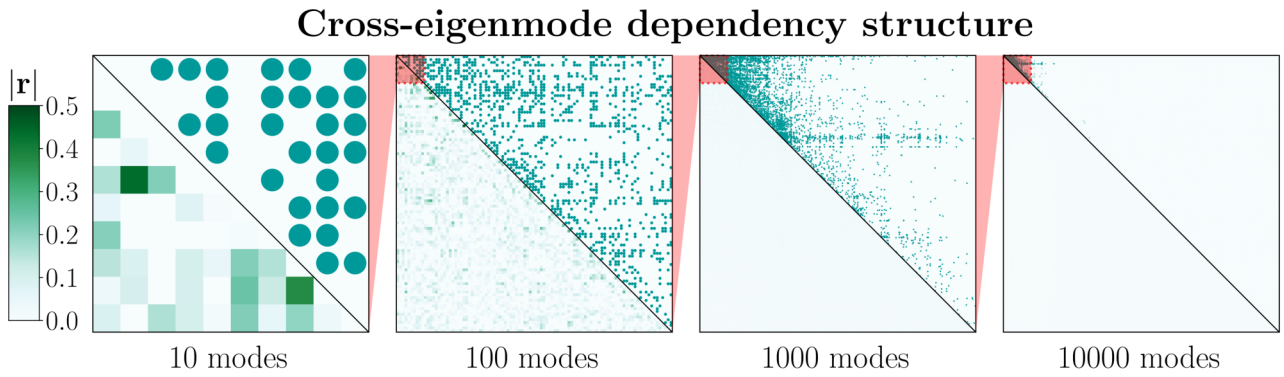
Each participant underwent neuroimaging and comprehensive clinical and neuropsychological evaluation, including the Mini-Mental State Examination (MMSE). MRI was performed on a 3T Siemens Magnetom Trio Tim scanner with a 32-channel head coil at the Clinical Imaging Research Centre of the National University of Singapore. The imaging protocol included a T1-weighted MPRAGE sequence (1 mm isotropic resolution, TR = 2300 ms, TE = 1.9 ms, TI = 900 ms, flip angle = 9°, 192 sagittal slices, matrix size = 256 × 256)[53]. T1-weighted scans were preprocessed using FreeSurfer 6.0 `recon-all` pipeline to generate cortical surfaces and thickness estimates in the fsnative surface space[54,55]. Matlab scripts and commands from Connectome Workbench and FreeSurfer were then used to project thickness data to the fs-LR surface space, aligning it with HCP thickness data to enable transfer learning from HCP datasets to MACC.

Cohort	**HC**	**MCI**	**AD**
Description	No cognitive impairment	Cognitive impairment no dementia	Diagnosis of AD
Sample Size	132	202	208
Age	68.49 ± 7.58	74.29 ± 6.70	75.49 ± 7.19
Sex	54.5% female	58.4% female	68.8% female
MMSE	27.39 ± 1.90	24.04 ± 3.88	15.60 ± 4.96

Table S.2. Sample Characteristic of MACC-H dataset. Note: Values presented as $\mu \pm \sigma$ indicate the mean (μ) and standard deviation (σ).

S.1.5 Cross-Basis Correlation Structure

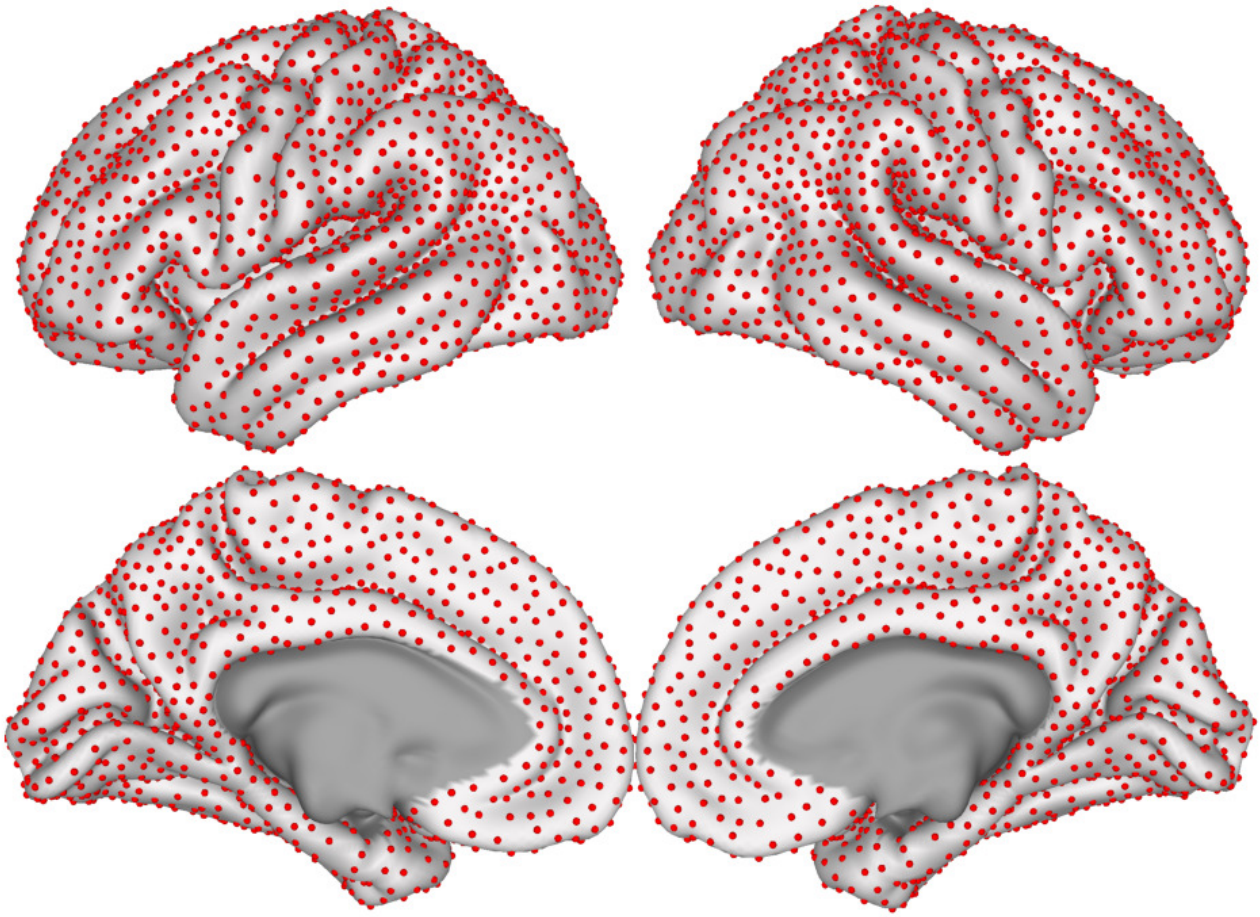
As described in Section 4.8.3 Spectral Normative Model (SNM), SNM leverages the naturally sparse correlation structure among spectral coefficients to efficiently approximate cross-mode dependencies. Subthreshold correlations are set to zero, reducing computational complexity while preserving key dependencies. Supplementary Fig. S.3 visualizes this sparsity by showing the magnitude of cross-basis correlations for cortical thickness phenotypes. A threshold was applied such that only the top 0.5% of all pairwise correlations were retained, producing a sparse matrix that captures the most relevant dependencies between eigenmodes.



Supplementary Fig. S.3. Sparse Cross-Basis Correlation Structure. Magnitudes of pairwise correlations between spectral coefficients encoding cortical thickness phenotypes. Nested heatmaps depict the sparsity of cross-mode dependencies: the lower triangle shows all correlation magnitudes, while the upper triangle highlights suprathreshold correlations retained by the model. The sparsity threshold corresponds to the top 0.5% of correlations, producing a sparse matrix that efficiently captures the most relevant cross-mode relationships.

S.1.6 Benchmarking High-Resolution Spatial Queries

Running direct models at full vertex resolution is computationally prohibitive, as a separate model must be fitted for each vertex. To enable a tractable vertex-wise benchmarking evaluation while preserving representative cortical coverage, we randomly sampled 4,000 vertices (2,000 per hemisphere). Left-hemisphere vertices were selected using a randomized distance-based k-means subsampling procedure ($k = 2{,}000$); the vertex-wise symmetry of the fs-LR atlas was then leveraged to include the homotopic right-hemisphere counterparts of these vertices, enabling assessment of asymmetric spatial queries (see Section S.1.8 Sensitivity Analyses for Query Characteristics). Supplementary Fig. S.4 shows the spatial distribution of the selected vertices. This subsample was used exclusively for benchmarking spectral normative model performance (Fig. 2); all other vertex-wise SNM analyses in the manuscript, including growth gradients and clinical applications, use the full vertex set, demonstrating the scalability of SNM to tens of thousands of concurrent spatial queries.



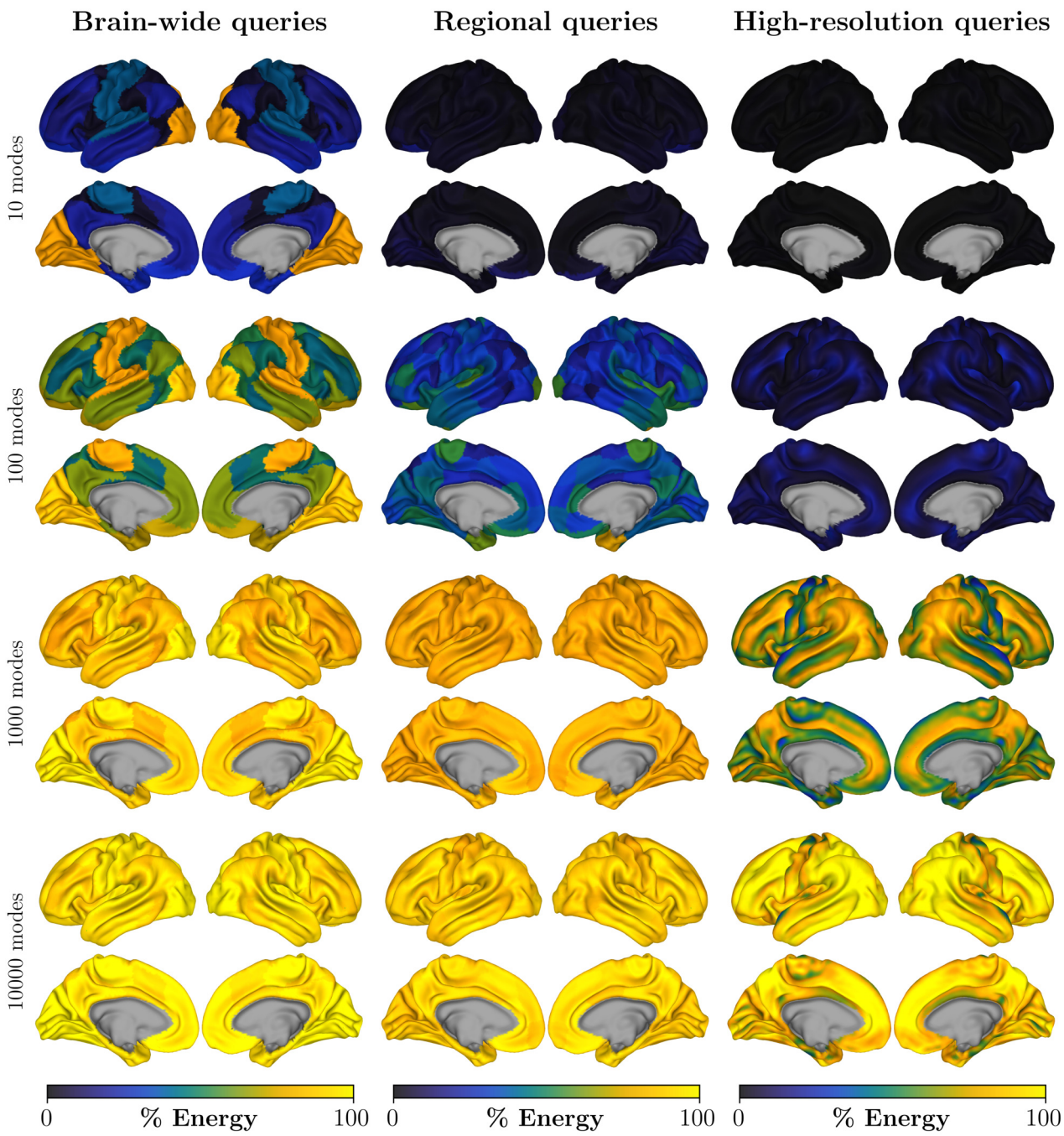
Supplementary Fig. S.4. Cortical Projection of Sampled High-Resolution Vertices. The 4,000 vertices used for benchmarking are indicated by red spheres on the fs-LR cortical surface. Vertices were selected using a distance-based subsampling procedure to ensure uniform coverage across the cortex. This selection was used only for benchmarking normative performance; all other SNM analyses in the manuscript use the full vertex set.

S.1.7 Cortical Projections of Performance Metrics

Extended Data Fig. E.1 presents a summary of reconstruction accuracy across different families of brain signals, and Fig. 2 summarizes comparative performance between SNM and the direct model. In the following sections, we expand upon these findings, detailing the regional distribution of performance metrics to offer deeper insights into comparative SNM evaluations.

S.1.7.1 Reconstruction Accuracy

Extended Data Fig. E.1 quantifies the reconstruction accuracy of different spatial query families in terms of the proportion of signal energy captured as a function of the number of modes. Fig. S.5 extends

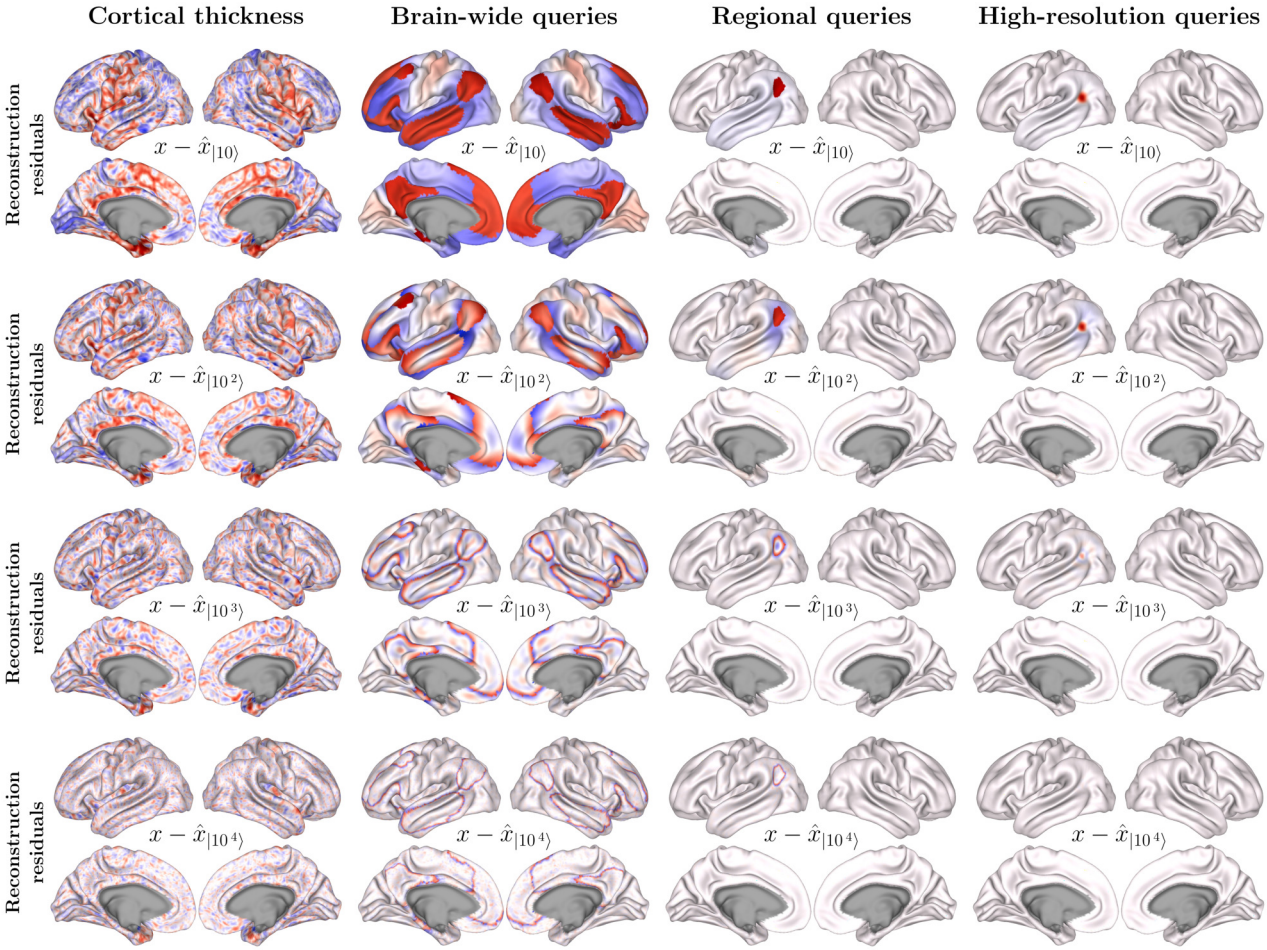


Supplementary Fig. S.5. Regional Variability of Reconstruction Accuracy across Spatial Query Families. Reconstruction accuracy is shown for brain-wide (Yeo 7 networks, left column), regional (Yan 200 atlas, middle column), and high-resolution (fs-LR 32k vertices, right column) spatial queries. Rows correspond to spectral reconstructions using $k = 10$, 10^2, 10^3, and 10^4 modes (top to bottom). For each query, reconstruction accuracy is quantified as the proportion of signal energy captured by the first k modes and projected onto the cortical location associated with the query.

this analysis by characterizing the regional variability of reconstruction accuracy for brain-wide (Yeo's 7 functional networks[56]), regional (Yan 200 atlas[57]), and high-resolution (fs-LR 32k vertices) spatial queries (see Section 4.9 Model Evaluation). For each query, the proportion of signal energy captured by the first $k \in \{10, 10^2, 10^3, 10^4\}$ modes was projected onto the corresponding cortical locations. These results further demonstrate that the first 10 modes retain only a small fraction of spatial variance across all query types, whereas approximately 100 modes begin to capture a considerable proportion of variance for brain-wide queries. In contrast, regional and high-resolution queries require at least 1,000 modes for accurate reconstruction, with further increases to 10,000 modes yielding diminishing incremental gains in reconstruction accuracy across all signal families.

S.1.7.2 Reconstruction Residuals

In Extended Data Fig. E.1, cortical projections of exemplary brain signals were presented alongside the original signal (last four rows). Supplementary Fig. S.6 shows the residuals of low-pass filtered graph approximations of the same signals. For four different low-pass approximations ($k = 10, 10^2, 10^3, 10^4$), the approximation residuals ($x - \hat{x}_{|k\rangle}$) are projected onto the cortical surface. As anticipated, the magnitude of the approximation error (indicated by color intensity in the cortical projections of residuals) decreases as the number of incorporated eigenmodes increases. Furthermore, with the inclusion of more modes, only regions containing high spatial frequency information (such as the transition loci along parcellation borders) have higher approximation errors. This verifies our expectations that low-pass filtering by eigenmodes can accurately approximate smooth signals (i.e. lower graph frequencies).

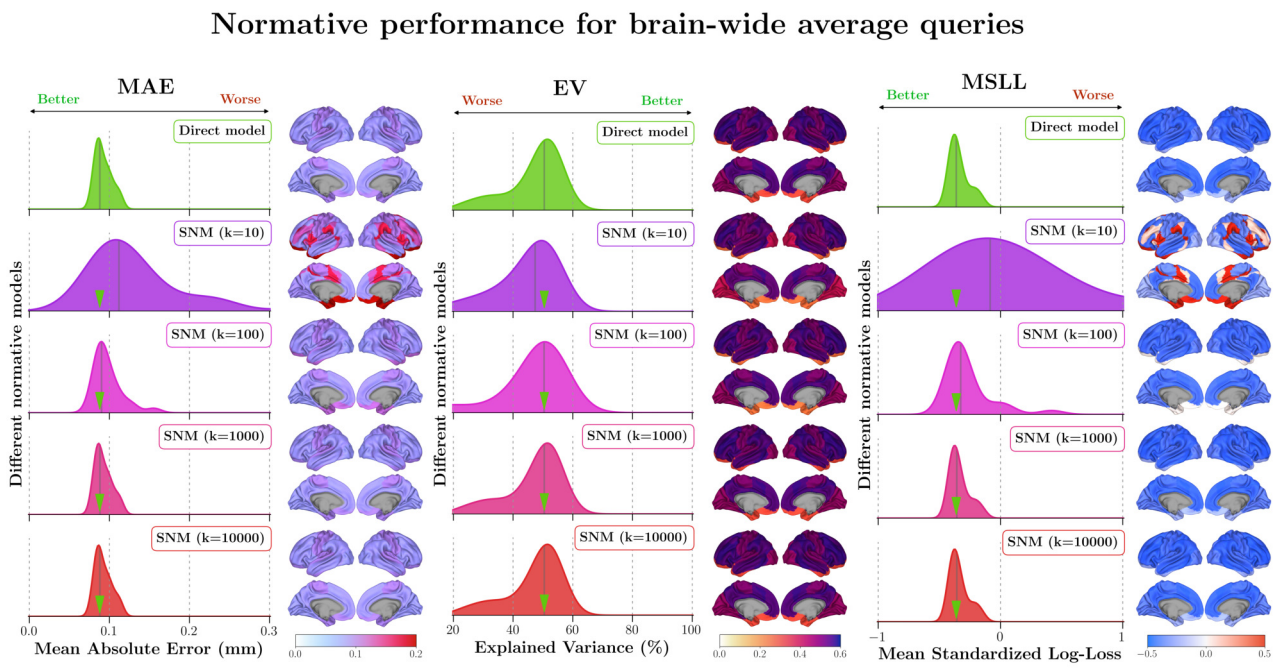


Supplementary Fig. S.6. Signal Reconstruction Residuals. The columns (from left to right) display the cortical projections of approximation residuals for individual cortical thickness signals, as well as brain-wide, regional, and high-resolution spatial queries (similar to Extended Data Fig. E.1). The rows show the residuals for the same signal reconstructed using 10, 100, 1,000, and 10,000 eigenmodes, respectively.

S.1.7.3 Brain-Wide Normative Estimates

Fig. S.7 presents the performance of various SNMs in reconstructing brain-wide thickness maps, compared to a direct model. The direct model is trained on observed cortical phenotypes for each spatial query ($y = T_{\text{train}} \,.\, x$), while SNMs rely solely on spectral approximations ($y \approx \sum_{i \leq k} s_i \tilde{x}_i$). Four SNMs are evaluated, which respectively utilize the first $k = \{10, 10^2, 10^3, 10^4\}$ eigenmodes.

With as few as 10 modes, spectral normative models can produce acceptable normative estimates for some brain-wide signals, indicated by negative MSLL values. However, specific brain-wide signals, such as the average thickness of the ventral attention network and the limbic network (from Yeo's 7 functional networks[56]), exhibit poor normative estimates at this level, showing noticeably inferior MAE and MSLL scores compared to the direct model. However, including at least 100 modes in the spectral normative model yields performance comparable to the direct model. These findings were

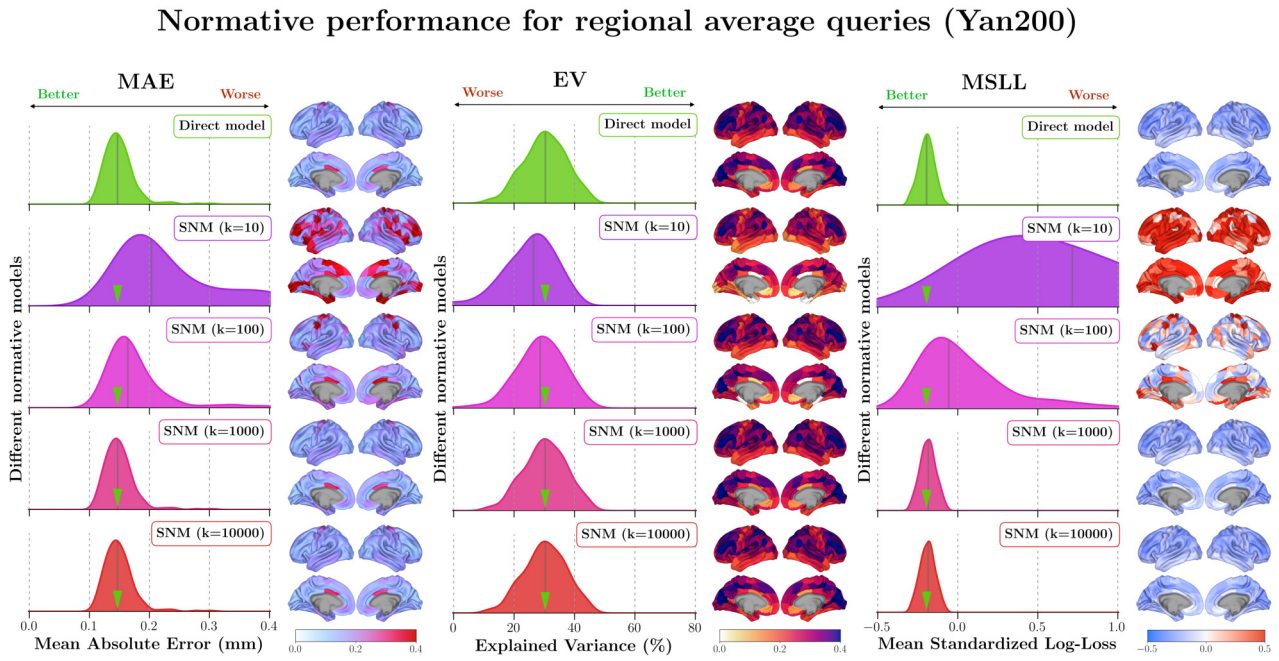


Supplementary Fig. S.7. Comparison of Performance for Different Normative Models Assessing Brain-Wide Spatial Queries. Normative performance is evaluated using out-of-sample assessments of direct and spectral normative models. The top row displays the performance of the direct model, while the next four rows show the performance of spectral models with 10, 100, 1,000, and 10,000 eigenmodes, respectively. Mean absolute error (MAE), explained variance (EV), and mean standardized log-loss (MSLL) are used to quantify model performance. Violin plots illustrate the distribution of performance measures across 25 different brain-wide signals, including cortex-wide average thickness and thickness averaged over 7 and 17 functional networks. Median values are indicated by gray lines on the violin plots. A green arrow marks the median performance of the direct model over spectral performance distributions for visual comparison. Cortical surface projections show normative performance values for 7 functional networks, with shared colorbars across all five rows to facilitate visual comparison.

also extended in sensitivity analysis for brain-wide asymmetric queries (see Supplementary Fig. S.14).

S.1.7.4 Regional Normative Estimates

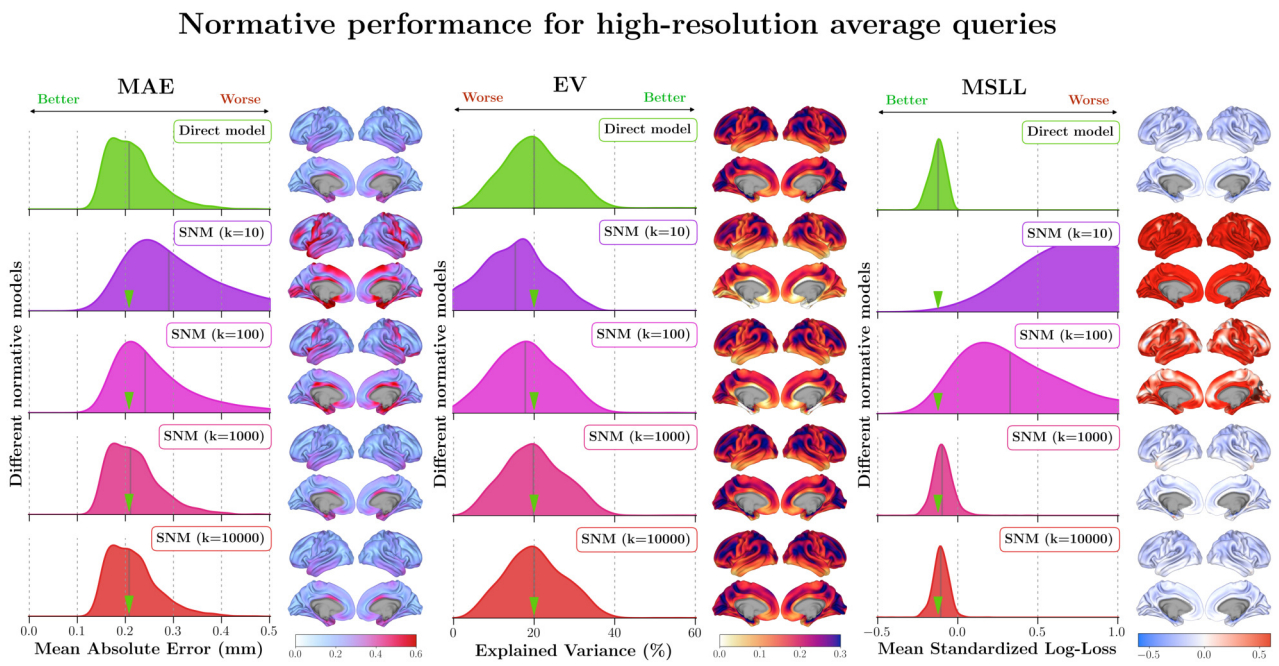
Fig. S.8 extends the previous comparison to regional signals. At this level of spatial specificity, SNM with only 10 modes shows substantially inferior performance across several brain regions. Including the first 100 modes results in acceptable performance for most brain regions, except for regions on the post-central gyrus, insular cortex, and orbitofrontal cortex. Incorporating up to 1000 modes in the spectral model achieves performance matching that of the direct model across all brain regions. Further inclusion of modes (up to 10,000) provides minimal additional benefits in goodness of fit, as measured by MAE, EV, and MSLL. Sensitivity analyses evaluate the robustness of these findings to regional granularity and signal asymmetry (see Supplementary Figs. S.11, S.15).



Supplementary Fig. S.8. Comparison of Performance for Different Normative Models Assessing Regional Spatial Queries. Normative performance is evaluated using out-of-sample assessments of both direct and spectral normative models. The top row shows the performance of the direct model, while the following four rows present the performance of spectral models with 10, 100, 1,000, and 10,000 eigenmodes, respectively. Mean absolute error (MAE), explained variance (EV), and mean standardized log-loss (MSLL) are used to quantify model performance. Violin plots display the distribution of performance measures across 200 different spatial queries, representing regional average thickness based on the Yan200 atlas. Median values are marked by gray lines on the violin plots. A green arrow marks the median performance of the direct model over spectral performance distributions for visual comparison. Cortical surface projections show normative performance values for each region, with shared colorbars across all five rows to aid visual comparison.

S.1.7.5 High-Resolution Normative Estimates

Finally, Fig. S.9 assesses SNM's performance in estimating high-resolution normative charts for spatial queries centered on various cortical surface voxels. Given the computational infeasibility of applying the direct model to every vertex, we evaluated a randomly selected subset of 400 vertices across the cortical surface. At this level of spatial specificity, spectral normative models with 10 and 100 modes fail to estimate normative ranges effectively compared to the direct model. However, the spectral model with 1000 modes achieves performance comparable to the direct model. Further inclusion of 10,000 modes yields marginal performance improvements. We also evaluated the sensitivity of these results to spatial granularity and asymmetry in normative queries (see Supplementary Figs. S.12, S.16).



Supplementary Fig. S.9. Comparison of Performance for Different Normative Models Assessing High-Resolution Spatial Queries. Normative performance is evaluated through out-of-sample assessments of both direct and spectral normative models. The top row displays the performance of the direct model, while the subsequent four rows show the performance of spectral models with 10, 100, 1,000, and 10,000 eigenmodes, respectively. Mean absolute error (MAE), explained variance (EV), and mean standardized log-loss (MSLL) are used to quantify model performance. Violin plots illustrate the distribution of performance measures across 4,000 different spatial queries, each representing an 8mm FWHM Gaussian kernel centered at a high-resolution cortical vertex (see Section S.1.6 Benchmarking High-Resolution Spatial Queries). Median values are indicated by gray lines on the violin plots. A green arrow highlights the median performance of the direct model in comparison to SNMs. Cortical surface projections show normative performance values for each vertex, depicted within its associated Yan400 region, with shared colorbars across all five rows to facilitate visual comparison.

S.1.8 Sensitivity Analyses for Query Characteristics

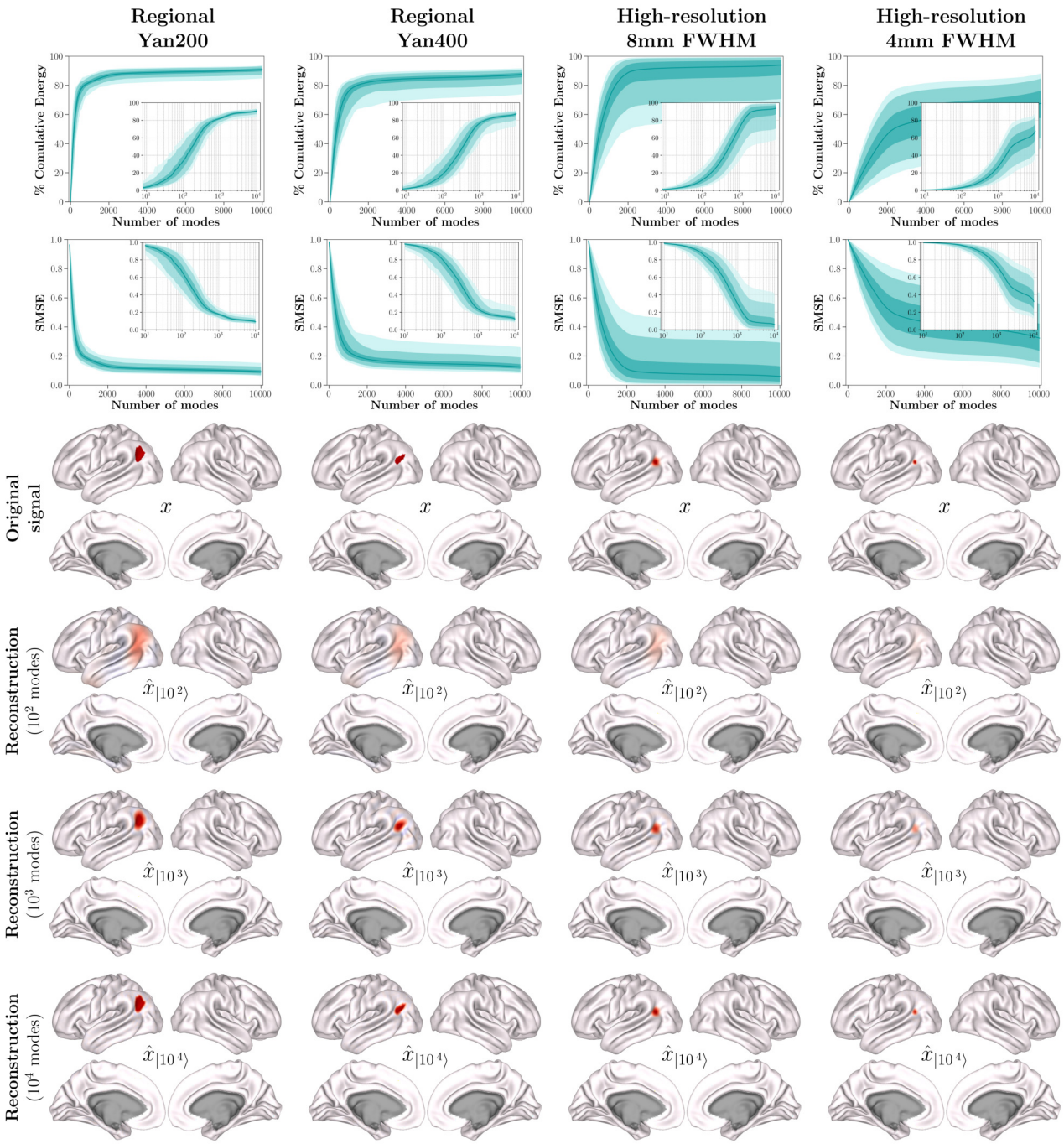
Several sensitivity analyses were conducted to evaluate the robustness of our findings to various spatial query properties. These analyses can be categorized into two main groups: those assessing the robustness of results against the granularity of the spatial queries, and those evaluating the impact of asymmetry in the spatial queries. These analyses will be respectively detailed in the ensuing sections.

S.1.8.1 Effect of Spatial Granularity

To evaluate the sensitivity of our findings to the spatial granularity of normative queries, we repeat the evaluations for both regional and high-resolution queries with altered parameters. For regional queries, we assess the impact of increased spatial granularity by using the Yan400 parcellation template, which contains twice as many brain regions as the Yan200 atlas, thereby doubling the spatial granularity of the regional signals. Similarly, the spatial granularity of the high-resolution signals are altered by adjusting the spatial smoothing kernel strength. Specifically, high-resolution evaluations are repeated using a 4mm FWHM Gaussian smoothing kernel, which effectively doubles the spatial granularity of the high-resolution signals.

Supplementary Fig. S.10 provides a comparative view of reconstruction accuracy performance as affected by signal granularity. As anticipated, increasing spatial granularity introduces higher spatial frequencies that are captured by higher graph frequencies. This is evident from the slight shift in proportional energy plots, which indicates an increased contribution of higher frequency modes to signal approximation. Additionally, the results show that the signal reconstruction accuracy for both regional and high-resolution spatial queries reaches a plateau after including the first 1,000 and 2,000 eigenmodes, respectively. Finally, the relatively higher values of SMSE for high-resolution signals with a 4mm FWHM smoothing kernel suggest that high-resolution spectral evaluations are better suited for signals smoothed to 8mm FWHM.

Supplementary Fig. S.11 illustrates the impact of regional signal granularity on the goodness of fit of normative models. Repeating the evaluations shown in Fig. S.8, we demonstrate that, irrespective

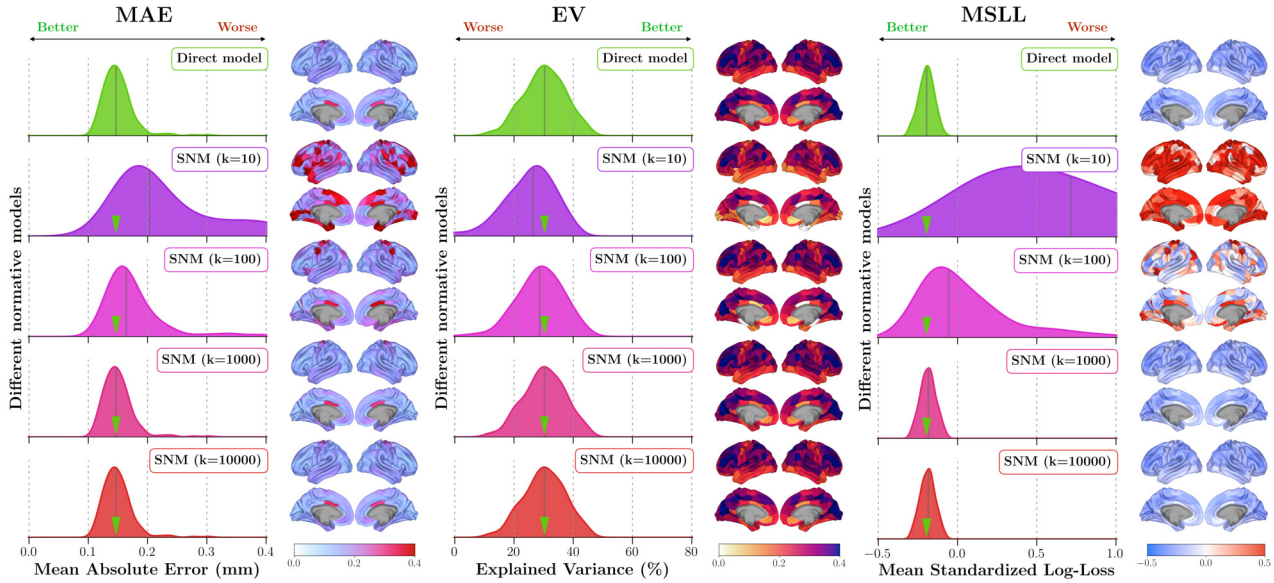


Supplementary Fig. S.10. Signal Reconstruction Accuracies across Different Granularity Scales. The evaluations presented in Extended Data Fig. E.1 are repeated to examine the effect of doubling the spatial granularity of the brain signals. The columns (from left to right) display spectral reconstruction accuracies for regional signals using the Yan200 and Yan400 atlases, as well as for high-resolution signals defined by 8mm FWHM and 4mm FWHM smoothing kernels. In the shaded line plots, the lines represent the median across all observations, while the shades indicate the [25, 75], [5, 95], and [1, 99] percentiles. The first row shows the proportional energy independently contributed by each eigenmode (logarithmic x-axis). The second row presents the standardized mean square error (SMSE) as a function of the number of low-frequency eigenmodes used for reconstruction (logarithmic x-axis for the insets). The third row illustrates one exemplary brain signal from each category, while the last three rows show the same signal reconstructed using 100, 1,000, and 10,000 eigenmodes, respectively.

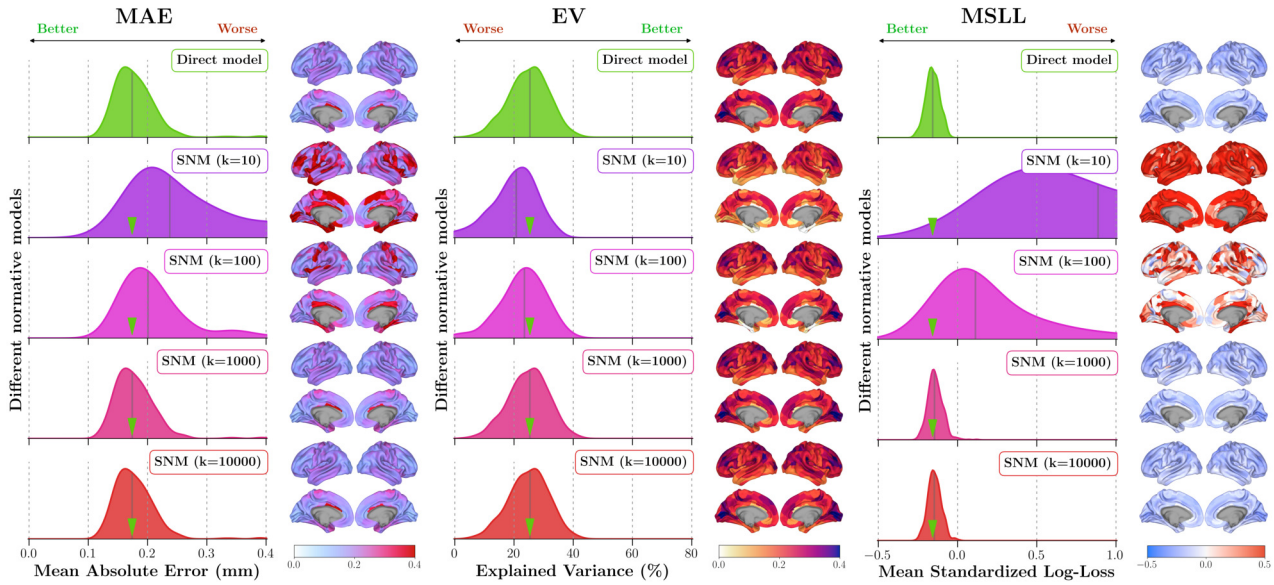
of the regional signal granularity, the inclusion of 1,000 eigenmodes in the spectral model achieves performance comparable to that of a direct model. Similarly, Supplementary Fig. S.12 repeats the evaluations presented in Fig. S.9 and shows that a spectral model with 1,000 modes yields normative estimates comparable to those produced by the direct model. While these results indicate the robustness of the spectral normative framework to changes in spatial granularity, any query used with the spectral model should nevertheless be tested to assess adherence to the low-pass spectral regime (e.g., by evaluating reconstruction accuracy as quantified by metrics such as SMSE).

Normative performance for regional queries: granularity

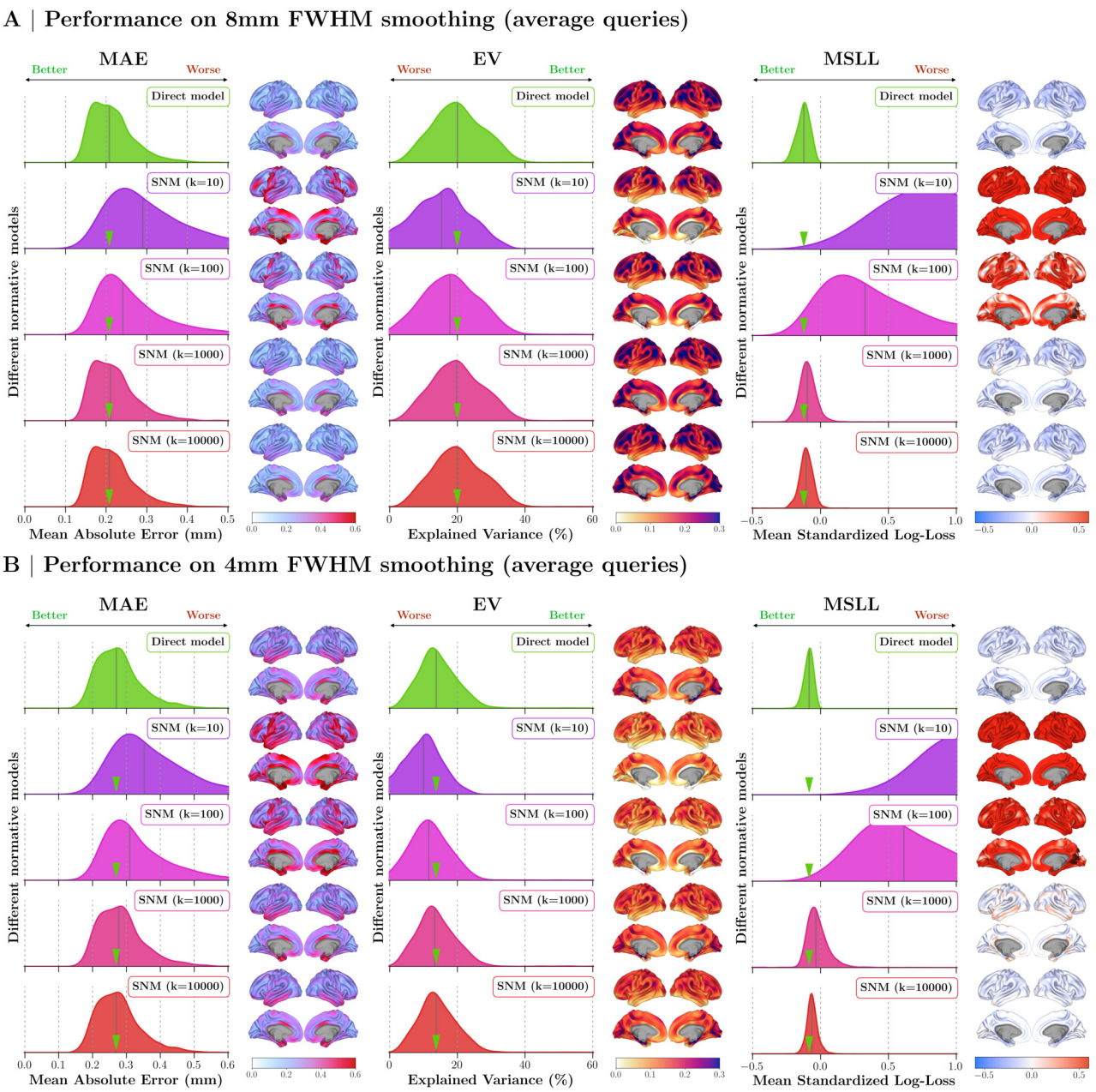
A | Performance on Yan200 atlas (average queries)



B | Performance on Yan400 atlas (average queries)



Supplementary Fig. S.11. Sensitivity of Normative Models Assessing Regional Spatial Queries at Different Granularity Scales. Normative performance is evaluated using out-of-sample assessments of both direct and spectral normative models. **(A)** same as Fig. S.8, repeated to aid visual comparison. **(B)** Same evaluations as panel A, but on regional queries from Yan400 atlas (increased spatial granularity).

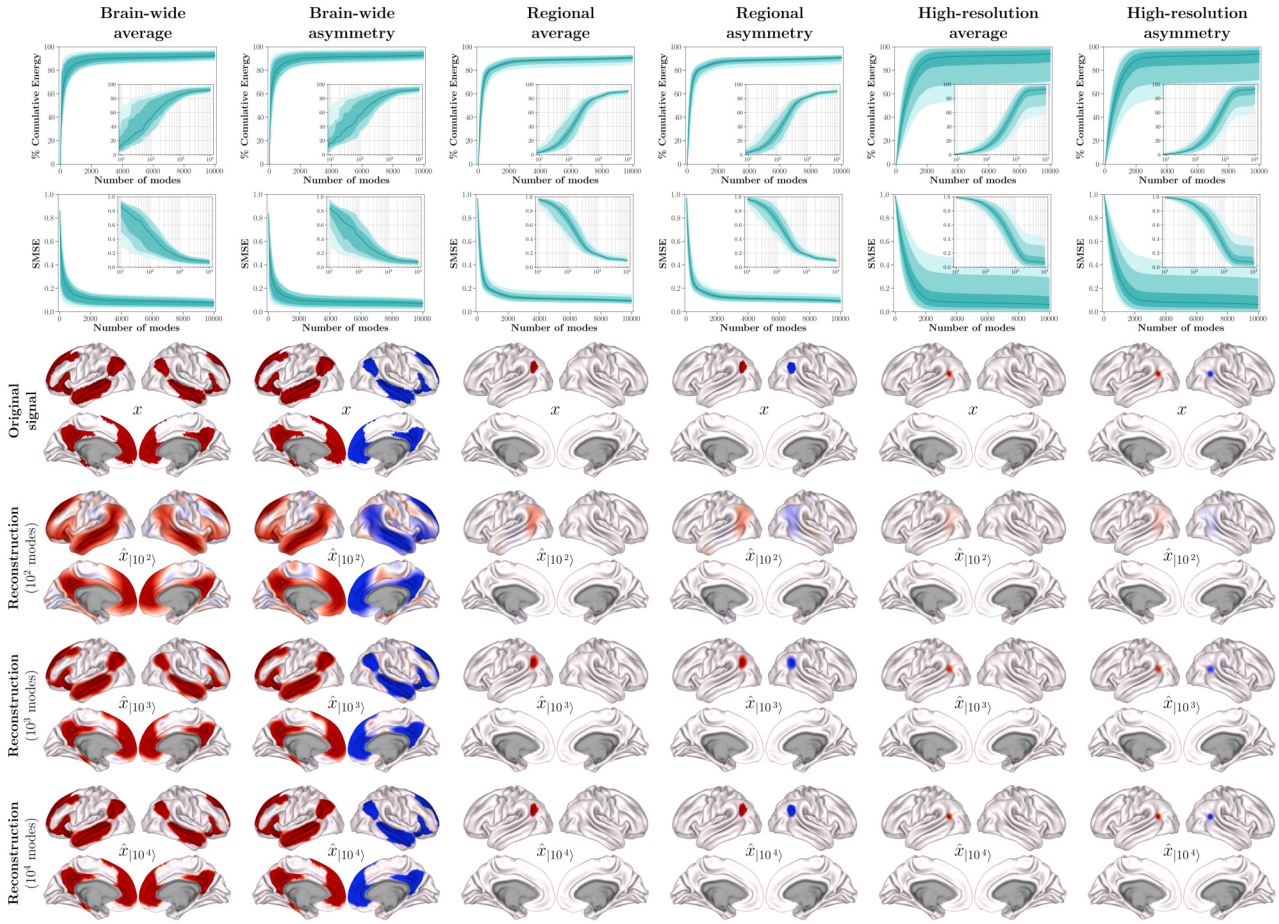


Supplementary Fig. S.12. Sensitivity of Normative Models Assessing High-Resolution Spatial Queries at Different Granularity Scales. Normative performance is evaluated using out-of-sample assessments of both direct and spectral normative models. **(A)** same as Fig. S.9, repeated to aid visual comparison. High-resolution spatial queries are generated using an 8mm FWHM Gaussian smoothing kernel. **(B)** Same evaluations as panel A, but on high-resolution queries generated using a 4mm FWHM Gaussian smoothing kernel (increased spatial granularity).

S.1.8.2 Effect of Spatial Asymmetry

A notable improvement of the spectral normative framework is its spatial versatility. This enables spectral models to extend beyond conventional average-based normative queries and gives them the ability to model a wider range of possible queries. As long as the query of interest adheres to the low-pass spectral regime, it can include various spatial patterns of interest. Asymmetric spatial queries are an interesting example of such patterns. Specifically, spectral normative models can be used to infer abnormal deviations in thickness lateralization by constructing a query that assesses the difference between the average thickness of a region/locus on the left hemisphere and its contralateral counterpart on the right hemisphere. This added benefit requires no extra training time and works out of the box. Hence, we evaluate the sensitivity of our main findings to the spatial asymmetry of normative queries.

We repeat the evaluations for brain-wide, regional, and high-resolution queries after introducing lateralization asymmetry. For brain-wide queries, we use a query that computes the average thickness

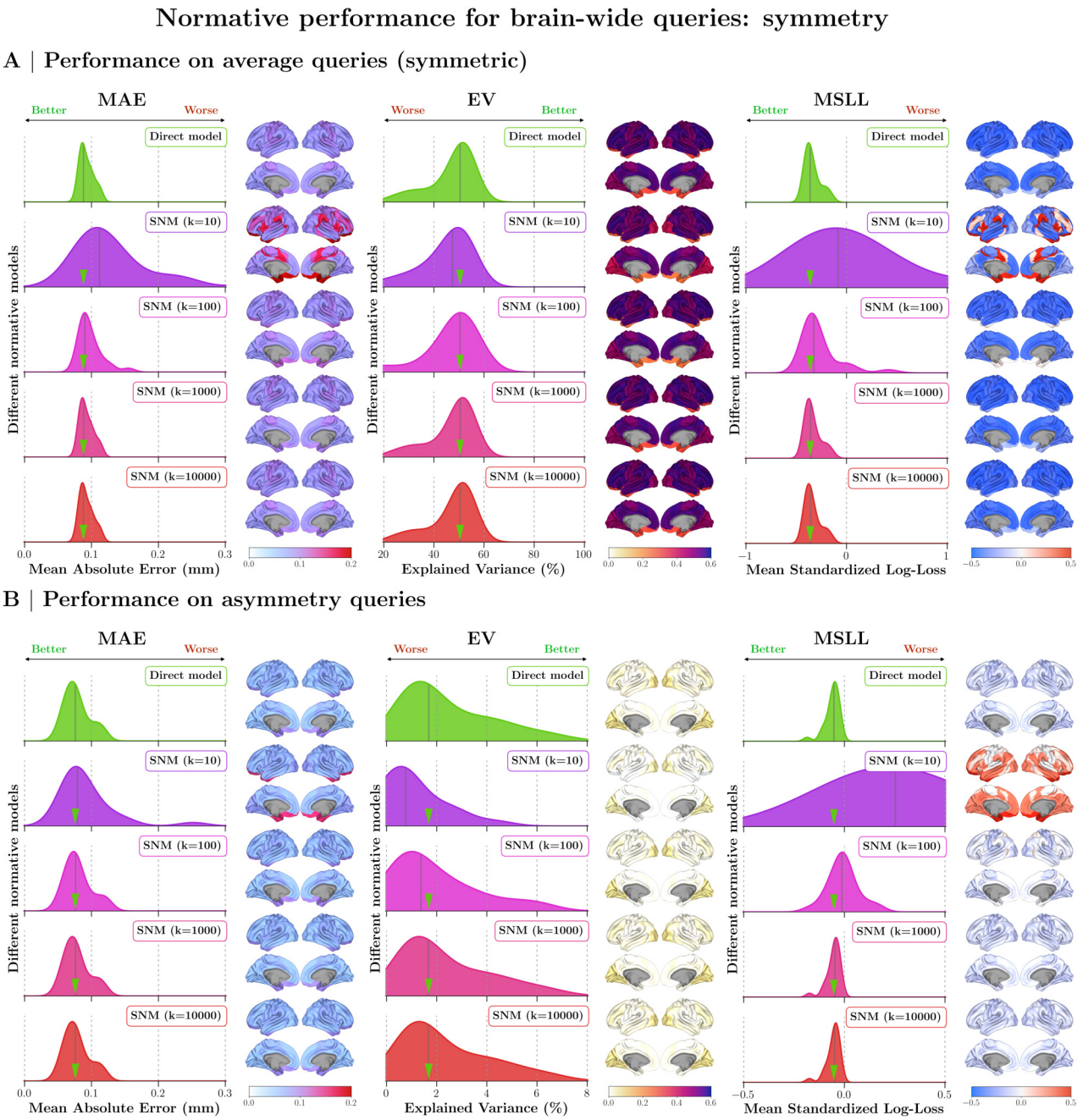


Supplementary Fig. S.13. Signal Reconstruction Accuracies for Average vs. Asymmetric Signals. The evaluations presented in Extended Data Fig. E.1 are repeated to examine the effect of introducing asymmetry to the brain signals. The columns display spectral reconstruction accuracies for pairs of (average, vs. lateralized/asymmetric) brain-wide, regional, and high-resolution signals. In the shaded line plots, the lines represent the median across all observations, while the shades indicate the [25, 75], [5, 95], and [1, 99] percentiles. The first row shows the proportional energy independently contributed by each eigenmode (logarithmic x-axis). The second row presents the standardized mean square error (SMSE) as a function of the number of low-frequency eigenmodes used for reconstruction (logarithmic x-axis for the insets). The third row illustrates one exemplary brain signal from each category, while the last three rows show the same signal reconstructed using 100, 1,000, and 10,000 eigenmodes, respectively.

of the query on the left cortex and subtracts it from the average of the part of the query that lies on the right cortex. For regional queries, given that the Yan atlas provides a homotopic parcellation of the cerebral cortex (i.e., parcels on the left and right cortices are paired), we evaluate 100 spatial normative queries, each assessing the difference in the average thickness of one homotopic parcel pair from Yan200 parcellation. Finally, for high-resolution queries, given that the fs-LR template space is aligned across the left and right cortices, we select pairs of homotopic vertices and construct spatial queries that compare their average thicknesses as described by 8mm FWHM smoothing kernels centered on respective vertex pairs.

Supplementary Fig. S.13 provides a comparative view of reconstruction accuracy performance as affected by signal asymmetry. Results indicate that signal asymmetry has negligible impact on eigenmode energy proportions and reconstruction accuracy (quantified by SMSE). Comparing asymmetric queries to average queries (presented in Extended Data Fig. E.1), we observe that asymmetric signals require a comparable number of modes for accurate reconstruction as average signals within each spatial query family (brain-wide, regional, and high-resolution). Regardless of signal symmetry, SMSE for the reconstruction of brain-wide, regional, and high-resolution signals respectively reaches below 0.2 after the inclusion of the first 400, 600, and 1,000 eigenmodes. Thus, the reported findings on the number of eigenmodes required to reconstruct brain signals remain consistent regardless of signal asymmetry.

Supplementary Figs. S.14, S.15, and S.16 illustrate the impact of signal asymmetry on the goodness of fit of the associated normative models. Repeating the evaluations shown in Figs. S.7, S.8, and S.9 shows that irrespective of the normative modeling framework (direct vs. spectral), goodness of fit,



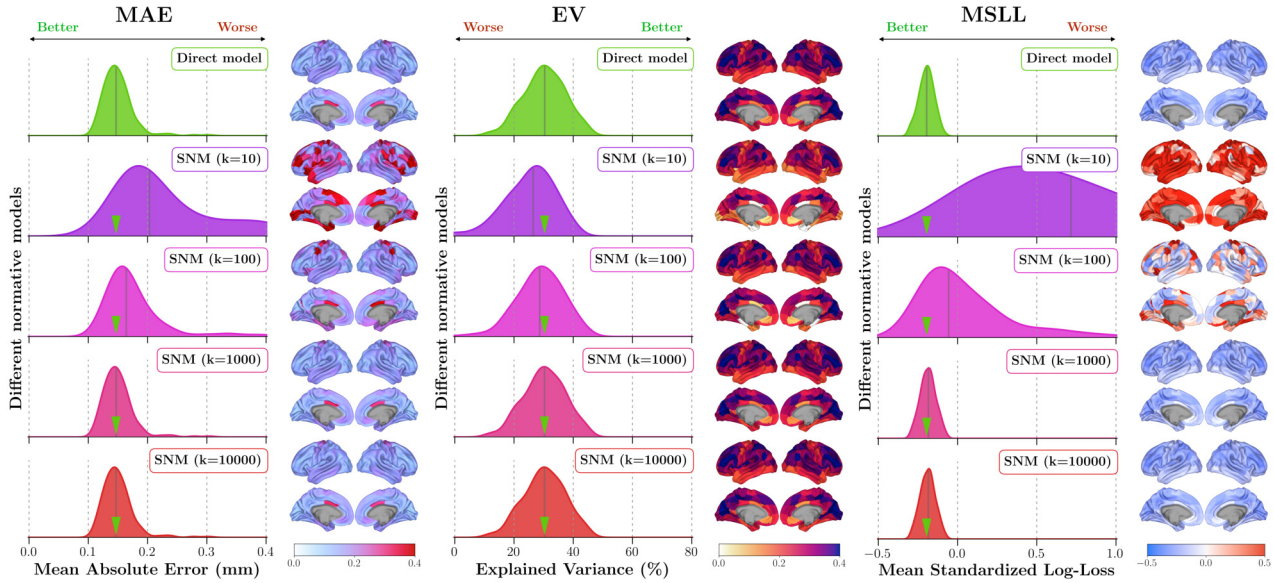
Supplementary Fig. S.14. Sensitivity of Normative Models to Assessing Lateralized Brain-Wide Spatial Queries. Normative performance is evaluated using out-of-sample assessments of both direct and spectral normative models. **(A)** same as Fig. S.7, repeated to aid visual comparison. **(B)** Same evaluations as panel A, but on lateralized counterparts of the brain-wide queries.

particularly when assessed by MSLL, is lower for lateralized norms. This effect is more pronounced at higher resolutions, suggesting that healthy norms of thickness lateralization are better studied at lower spatial specificity (functional networks or parcels). As the goodness of fit in predicting the central tendency (quantified via MAE) is less affected, we speculate that this reduction in the accuracy of lateralized queries is due to the misaligned gyrification patterns between contralaterally aligned high-resolution vertices, rendering lateralization evaluations less meaningful at higher resolutions.

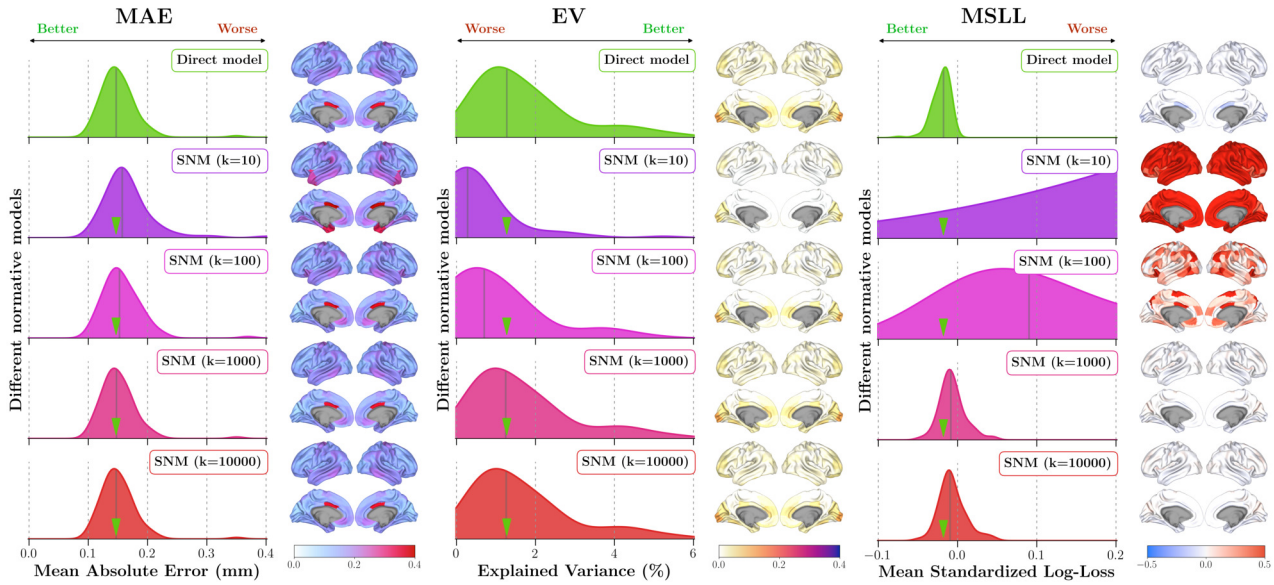
Despite this effect, which was not specific to the spectral model, we observe that, similar to the main findings, the inclusion of 100 eigenmodes is sufficient to model brain-wide normative thickness ranges with the spectral model, achieving performance on par with the direct model. As with the main findings, the number of modes required to achieve comparable performance to that of a direct model at the resolution of regional or vertex-wise queries increases to 1,000 eigenmodes. These sensitivity evaluations demonstrate that our findings are robustly replicable for signals with asymmetry, and that the inclusion of 1,000 eigenmodes can provide comparable performance to that of a conventional model across a wide range of spatial queries.

Normative performance for regional queries: symmetry

A | Performance on average queries (symmetric)



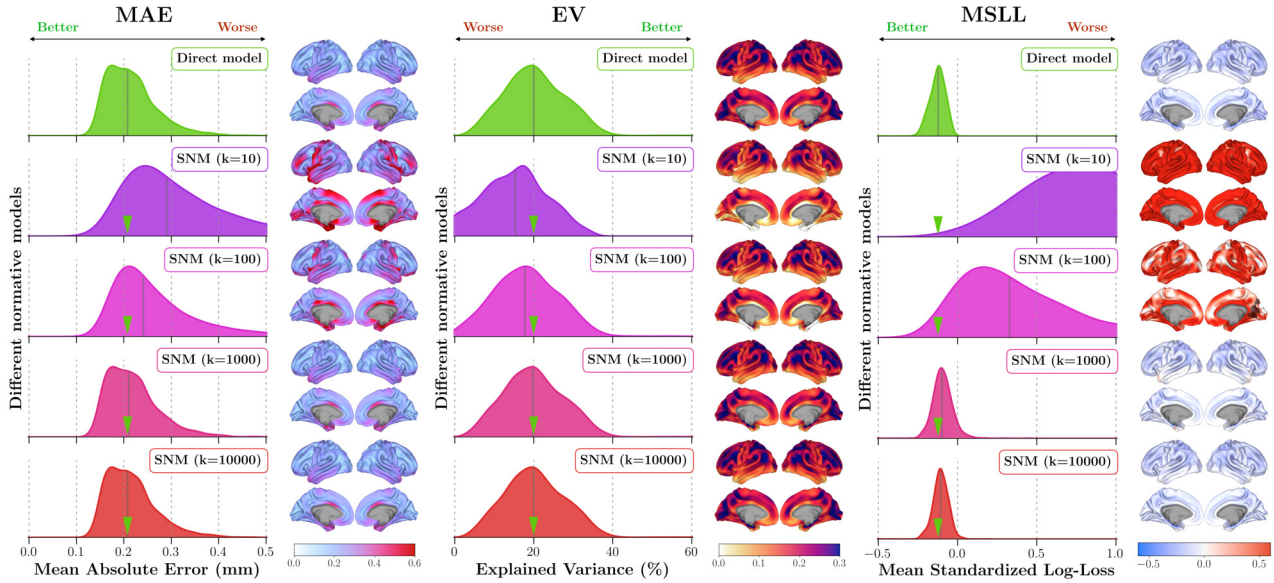
B | Performance on asymmetry queries



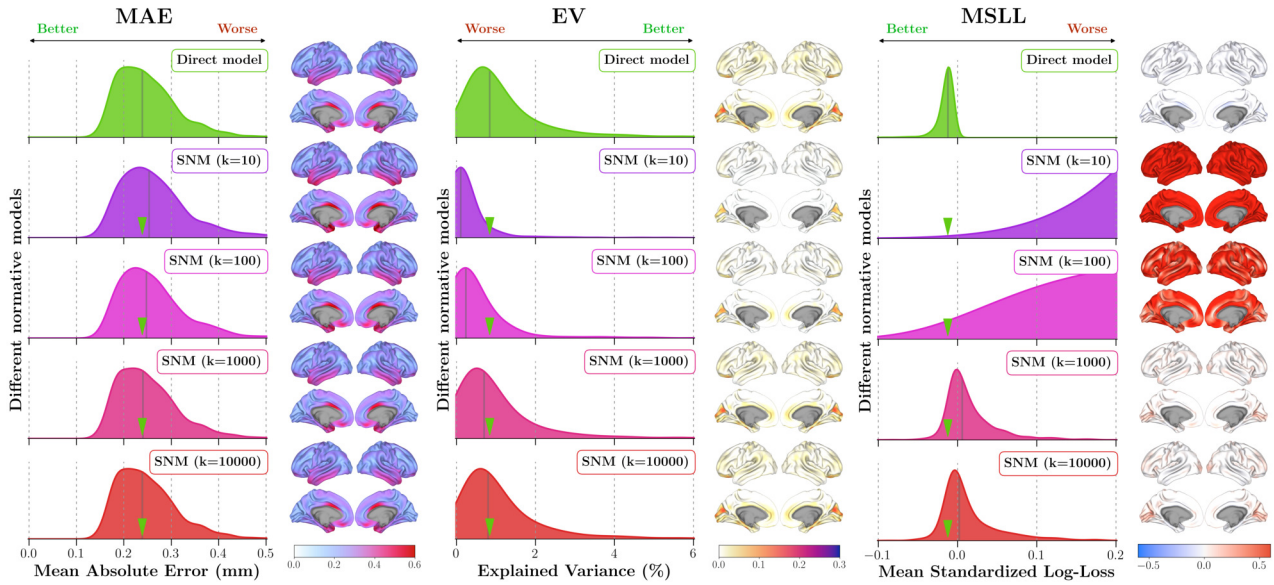
Supplementary Fig. S.15. Sensitivity of Normative Models to Assessing Lateralized Regional Spatial Queries. Normative performance is evaluated using out-of-sample assessments of both direct and spectral normative models. **(A)** same as Fig. S.8, repeated to aid visual comparison. **(B)** Same evaluations as panel A, but on lateralized counterparts of the regional queries.

Normative performance for high-resolution queries: symmetry

A | Performance on average queries (symmetric)



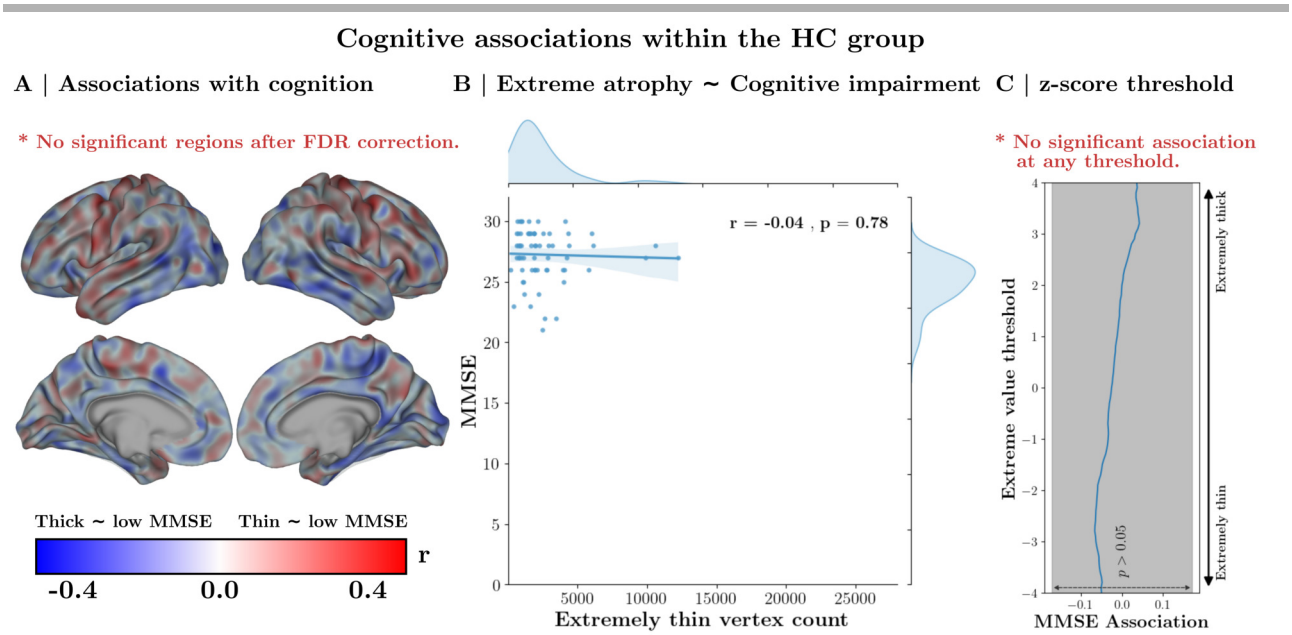
B | Performance on asymmetry queries



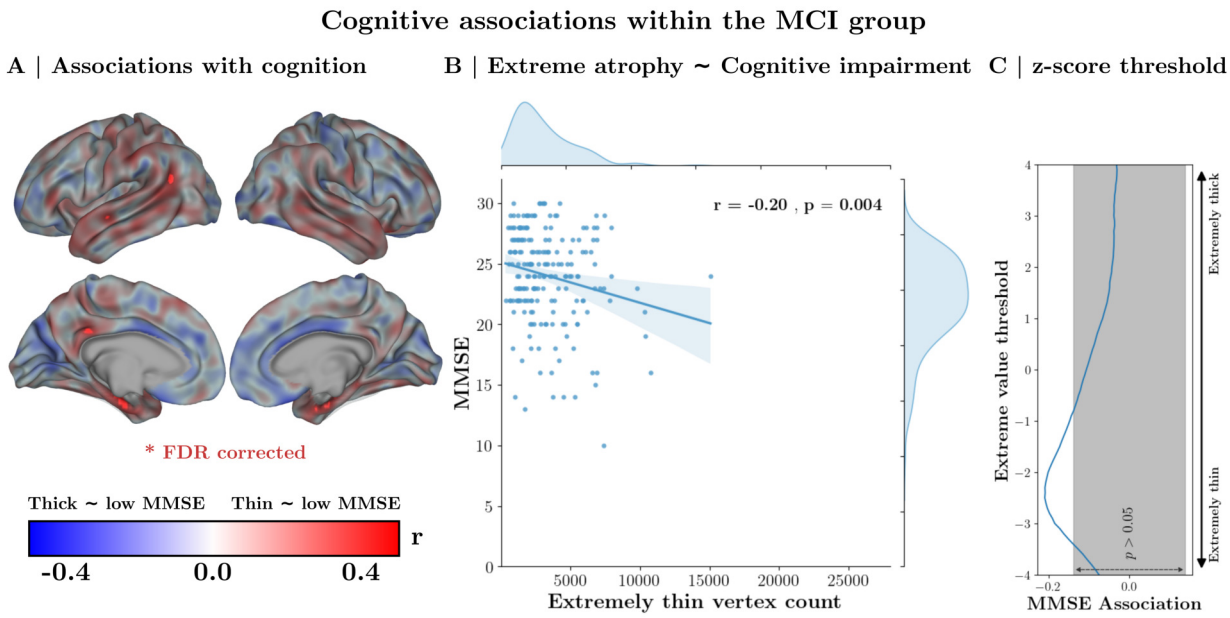
Supplementary Fig. S.16. Sensitivity of Normative Models to Assessing Lateralized High-Resolution Spatial Queries. Normative performance is evaluated using out-of-sample assessments of both direct and spectral normative models. **(A)** same as Fig. S.9, repeated to aid visual comparison. **(B)** Same evaluations as panel A, but on lateralized counterparts of the high-resolution queries.

S.1.9 Within-Group Cognitive Associations

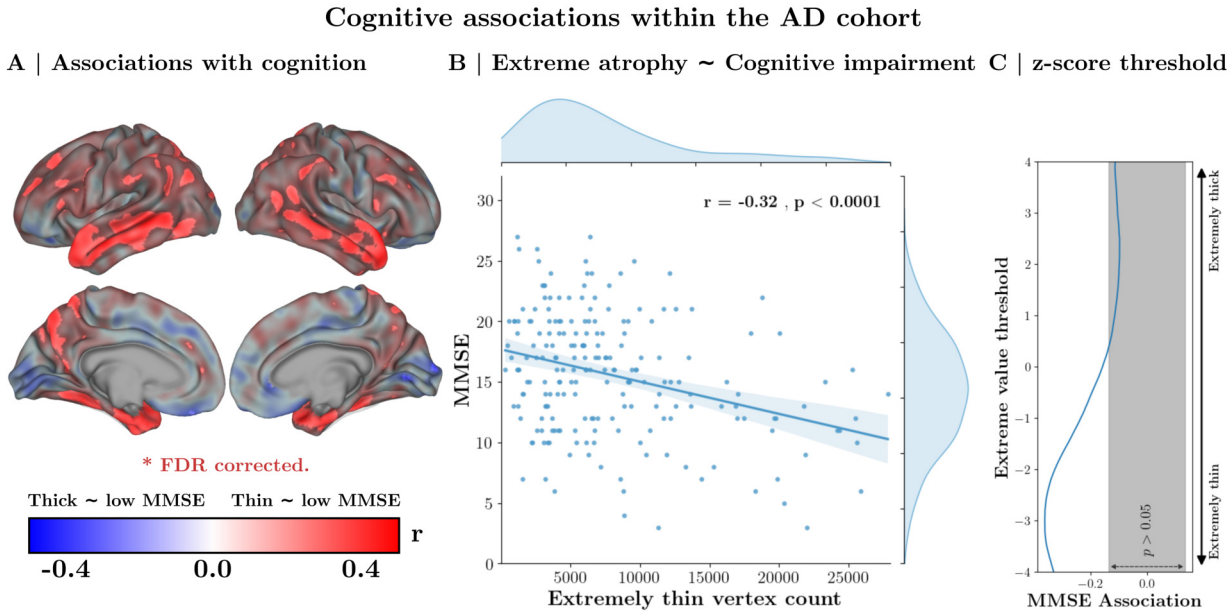
To assess the significance of cognitive associations within each cohort, we repeated the analyses from Section 2.5 SNM Uncovers Cortical Signatures of Atrophy in Alzheimer's Disease on the HC, MCI, and AD subsets, with the expectation of no association in HC, weaker associations in MCI, and stronger effects in AD. For healthy controls (HC), high-resolution cortical associations did not reach significance in any region after FDR correction (Supplementary Fig. S.17), and ETVC showed no association with cognitive performance ($r = -0.04$, $p = 0.78$), irrespective of the z-threshold. In the MCI cohort, ETVC showed a relatively weak but significant association with cognitive impairment ($r = -0.20$, $p = 0.004$; Supplementary Fig. S.18), while regional analyses identified atrophy in the entorhinal cortex to be associated with MCI cognitive impairment. In contrast, the AD cohort displayed widespread neocortical atrophy patterns significantly linked to cognitive impairment (Supplementary Fig. S.19), with reductions in the temporal pole, superior and middle temporal gyri, entorhinal cortex, parahippocampal cortex, precuneus, supramarginal gyrus, and superior, middle, and inferior frontal gyri. Additionally, ETVC showed a stronger association with cognitive impairment in the AD cohort ($r = -0.32$, $p < 0.0001$). Interestingly, this predictive capacity diminished when the z-score threshold was raised to zero or any positive value, with optimal predictions achieved between z-score thresholds of -2 and -3. Together, these findings suggest that the derived normative atrophy marker of cognitive impairment exhibits increasing sensitivity from MCI to AD and is most informative in the presence of extreme atrophy.



Supplementary Fig. S.17. Cognitive Association Tests within the Healthy Cohort (HC). (A) Vertex-level normative assessments were tested for linear associations with cognitive performance (MMSE), but no vertices reached significance after nonparametric FDR correction. **(B)** Extreme value statistics for severe atrophy (ETVC, z-score threshold of -1.96) were assessed for their predictive capability on cognitive performance, revealing no sensitivity of ETVC to cognitive variations in HC. **(C)** A range of alternative z-score thresholds was tested, demonstrating that ETVC does not predict cognitive performance in HC, regardless of threshold choice.



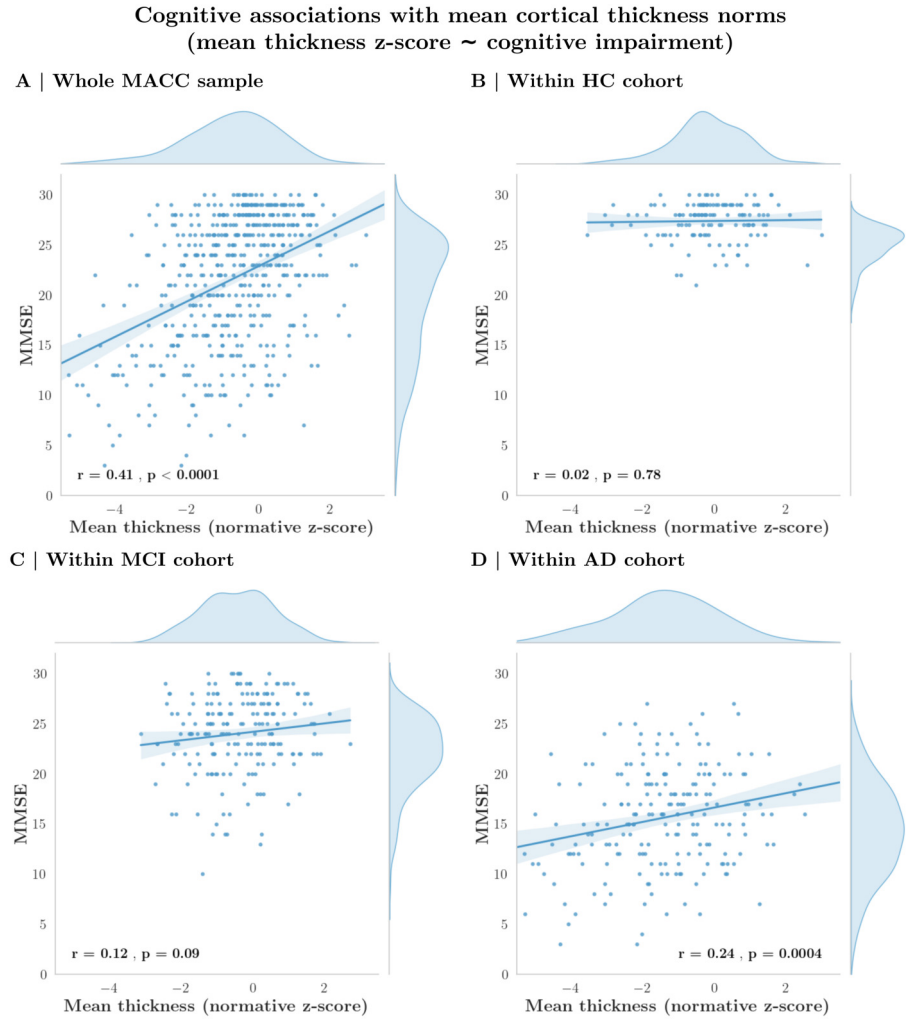
Supplementary Fig. S.18. Cognitive Association Tests within the MCI Cohort. (A) Vertex-level normative assessments were tested for linear associations with cognitive performance (MMSE) within the MCI cohort. A limited set of vertices reached statistical significance, indicating that extreme atrophy in specific regions may serve as a marker of cognitive impairment. **(B)** Extreme value statistics for severe atrophy (ETVC, z-score threshold of -1.96) were evaluated for their predictive capability on cognitive performance, showing a modest association between ETVC and cognitive variations within MCI ($r = -0.20$). **(C)** Testing a range of alternative z-score thresholds shows that extreme focal atrophy ($-3 < z < -2$) can predict cognitive performance in MCI.



Supplementary Fig. S.19. Cognitive Association Tests within the AD Cohort. (A) Vertex-level normative assessments were tested for linear associations with cognitive performance (MMSE) within the AD cohort. A widespread set of cortical regions reached statistical significance, highlighting extensive brain-wide associations as markers of AD-related cognitive impairment. **(B)** Extreme value statistics for severe atrophy (ETVC, z-score threshold of -1.96) were evaluated for predictive capability regarding cognitive performance, revealing significant predictive power of ETVC to capture cognitive variations within the AD cohort. **(C)** Testing across a range of alternative z-score thresholds indicated that only ETVCs with a negative z-threshold significantly predicted cognitive performance in the AD cohort. Notably, extreme thinning (z-thresholds from -2 to -3) proved to be the strongest predictor of cognitive impairment.

S.1.10 Brain-Wide Cognitive Associations

To highlight the benefits of inferring high-resolution spatial norms, we compare the predictive performance of the ETVC normative measure derived from high-resolution brain charting with that of a brain-wide normative z-score computed from a spatially coarse brain chart of mean cortical thickness (averaged across the entire cortex). Using the fine-tuned SNM, we calculate individual z-scores for deviations in average cortical thickness and replicate the evaluations from Figs. 5C, S.17B, S.18B, and S.19B. As shown in Fig. S.20, the results consistently indicate that while brain-wide normative measures can predict cognitive impairments, high-resolution ETVC measures outperform them. Specifically, ETVC achieves higher prediction accuracy in the entire sample (ETVC: $|r| = 0.53$, brain-wide: $|r| = 0.41$) and in the AD subsample (ETVC: $|r| = 0.32$, brain-wide: $|r| = 0.24$).



Supplementary Fig. S.20. Cognitive Association Tests with a Coarse Brain-Wide Normative Metric. Normative deviations (z-scores) in mean cortical thickness were evaluated for their predictive capability regarding cognitive performance (quantified by MMSE). Predictive performance was assessed in **(A)** the entire MACC dataset, as well as within the **(B)** HC, **(C)** MCI, and **(D)** AD cohorts individually. While the global association is robust, the attenuated correlation in HC and MCI is attributed to restricted variance and a ceiling effect in MMSE scores. Sensitivity evaluations in Section S.1.11 Association Consistency Across Diagnostic Cohorts confirm that the direction of effect is preserved across groups.

S.1.11 Association Consistency Across Diagnostic Cohorts

The evaluation of brain-behavior associations across a heterogeneous clinical spectrum necessitates a verification that observed global trends are representative of underlying within-group relationships. Specifically, if diagnostic cohorts occupy distinct ranges of both the predictor (biomarker) and the outcome (MMSE), a pooled correlation could potentially be driven by mean differences between groups rather than a consistent biological effect. To ensure the robustness of the associations presented in Sections S.1.10 and S.1.9, we performed a series of sensitivity analyses to test for slope consistency and to control for group-level baseline shifts.

First, we implemented Linear Mixed-Effects Models (LMM) for both biomarkers (ETVC and mean thickness z-score), treating the diagnostic group as a random intercept to account for group-specific baseline shifts in MMSE (MMSE $\sim$ Biomarker+(1|Diagnosis)). As detailed in Table S.3 and Table S.4, the fixed effect of the biomarker remained highly significant for both ETVC ($z = -6.53, p < 0.001$) and mean thickness z-scores ($z = 4.28, p < 0.001$). This confirms that the brain-behavior associations

persist and remain significant after explicitly accounting for the variance explained by diagnostic groups (HC, MCI, and AD).

Second, we performed interaction modeling (MMSE $\sim$ Biomarker $\times$ Diagnosis) to formally test for slope heterogeneity between cohorts. For ETVC, no significant interactions were found (all $p > 0.17$, see Table S.5), indicating that the relationship between high-resolution normative deviations and cognitive impairment is statistically consistent across the entire clinical spectrum. For the mean thickness z-score, we observed an attenuation of the slope specifically in the HC group ($p = 0.052$ for the interaction term, Table S.6). However, as visualized in Fig. S.20B, this lack of significant correlation in healthy individuals is attributable to a marked ceiling effect and restricted variance in MMSE scores, rather than a reversal or absence of the underlying biological trend. Collectively, these results demonstrate that the direction of effect is preserved across the clinical spectrum and that the global associations reported in this study are robust to group-level confounds.

Variable	**Coef.**	**Std.Err.**	**z**	**P> |z|**	**[0.025, 0.975]**
Fixed Effects					
Intercept	23.5	3.1	7.645	0.000	[17.5, 29.6]
ETVC	-2.6e-04	4.0e-05	-6.532	0.000	[-3.4e-04, -1.8e-04]
Random Effects	Variance (Std.Err.)				
Group Variance	28.2 (7.4)				

Table S.3. Mixed Linear Model Results for ETVC. The model (MMSE $\sim$ ETVC + (1|Diagnosis)) treats diagnostic groups as random intercepts. The highly significant z-statistic for ETVC confirms the association persists after accounting for group-level baseline shifts.

Variable	**Coef.**	**Std.Err.**	**z**	**P> |z|**	**[0.025, 0.975]**
Fixed Effects					
Intercept	22.7	3.3	6.891	0.000	[16.2, 29.1]
Mean Thickness (z)	0.5	0.1	4.285	0.000	[0.3, 0.8]
Random Effects	Variance (Std.Err.)				
Group Variance	32.4 (8.3)				

Table S.4. Mixed Linear Model Results for Mean Thickness z-score. The model (MMSE $\sim$ Mean Thickness z + (1|Diagnosis)) confirms that average cortical thickness remains a significant predictor of cognitive performance even when accounting for diagnostic group as a random effect.

Variable	**Coef.**	**Std.Err.**	**t**	**P> |t|**	**[0.025, 0.975]**
Main Effects					
Intercept (AD)	17.7	0.4	40.478	0.000	[16.8, 18.5]
Diagnosis [MCI]	7.5	0.6	11.606	0.000	[6.2, 8.7]
Diagnosis [HC]	9.8	0.7	14.796	0.000	[8.5, 11.1]
ETVC	-2.7e-04	4.4e-05	-6.019	0.000	[-3.5e-04, -1.8e-04]
Interaction Terms					
ETVC $\times$ MCI	-7.0e-05	1.3e-04	-0.554	0.580	[-3.2e-04, 1.8e-04]
ETVC $\times$ HC	2.2e-04	1.6e-04	1.349	0.178	[-9.8e-05, 5.3e-04]

Table S.5. Interaction Analysis for ETVC. The OLS model (MMSE $\sim$ ETVC $\times$ Diagnosis) tests for slope differences across groups. Interaction terms ($p > 0.17$) are non-significant, indicating that the relationship between ETVC and MMSE is consistent across diagnostic categories.

Variable	Coef.	Std.Err.	t	P> |t|	[0.025, 0.975]
Main Effects					
Intercept (AD)	16.6	0.4	46.266	0.000	[15.9, 17.3]
Diagnosis [MCI]	7.6	0.5	16.359	0.000	[6.7, 8.5]
Diagnosis [HC]	10.8	0.5	21.662	0.000	[9.8, 11.7]
Mean Thickness (z)	0.7	0.2	4.393	0.000	[0.4, 1.0]
Interaction Terms					
Mean Thickness (z) × MCI	-0.3	0.3	-0.998	0.319	[-0.9, 0.3]
Mean Thickness (z) × HC	-0.7	0.3	-1.946	0.052	[-1.4, 0.0]

Table S.6. Interaction Analysis for Mean Thickness z-score. The OLS model (MMSE ∼ Mean Thickness z × Diagnosis) explores slope variations across cohorts. While an attenuation of the slope is observed in the HC group ($p = 0.052$), the overall direction of the effect is consistent across groups, supporting the biological continuum.

S.1.12 Heterogeneity Landscape Subgroups

Fig. 6D presents snapshots of cortical deviation patterns across various regions within the 2-dimensional atrophy landscape. To more comprehensively illustrate the overall cortical signature in this landscape, we have provided four supplementary videos (one each for the whole sample, HC, MCI, and AD subsets). These videos display changes in the cortical deviation patterns as the local region is gradually shifted over a circular area within the 2-dimensional space.

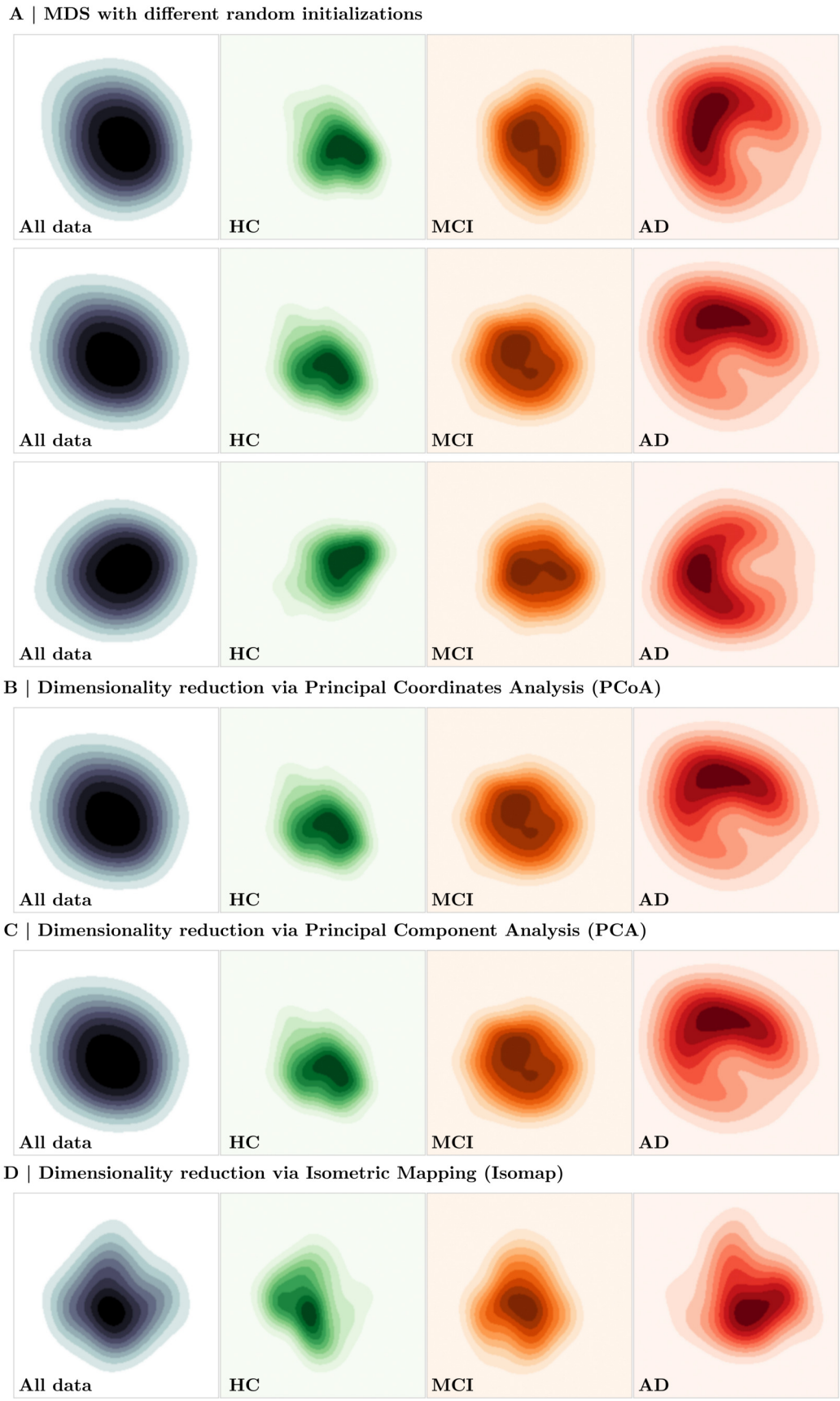
S.1.13 Robustness of the Heterogeneity Landscape

To assess the robustness of the heterogeneity landscape reported in the main text (Fig. 6), we performed a series of sensitivity analyses examining whether the reported qualitative characteristics depend on the choice of dimensionality reduction algorithm or associated hyperparameters. These analyses were designed to test whether the observed structure reflects stable properties of the underlying high-dimensional deviation space, rather than artifacts of a specific dimensionality reduction procedure.

First, using the same multidimensional scaling (MDS) framework adopted in the main analysis, we repeated the embedding under 3 alternative random seed initializations (Fig. S.21A). Across all runs, the principal visual features described in the Results were consistently reproduced. Specifically, healthy individuals formed a dense central cluster, MCI participants exhibited increased dispersion relative to controls, and AD-diagnosed individuals showed the greatest separation from the healthy cluster, with pronounced heterogeneity in the direction of deviations across individuals. These characteristics were preserved irrespective of seed choice, indicating that the observed structure is not sensitive to this hyperparameter.

Second, to evaluate whether similar landscape organization emerges under different dimensionality reduction paradigms, we repeated the analysis using principal coordinates analysis (PCoA, Fig. S.21B), principal component analysis (PCA, Fig. S.21C), and Isomap (Fig. S.21D). While these methods differ in their geometric assumptions and may yield embeddings with different orientations or relative scaling, all approaches consistently reproduced the core qualitative patterns observed with MDS. In particular, each method revealed a central concentration of healthy individuals, progressive dispersion in MCI, and a highly heterogeneous AD distribution spanning multiple directions in the latent space, without evidence of clear boundaries between potential subtypes.

Taken together with the statistical analyses reported in the Results, including group differences in mean pairwise distances and Hartigan's dip tests for clusterability[58,59], these sensitivity analyses indicate that the heterogeneity landscape reflects stable structure present in the high-dimensional normative deviation space. The observed patterns are therefore not specific to any single dimensionality reduction method, but instead represent an emergent property that is both visually apparent across embeddings and quantitatively supported by the underlying distance metrics.



Supplementary Fig. S.21. Robustness of the Heterogeneity Landscape to Dimensionality Reduction Choices. Kernel density estimates (KDEs) of the 2D heterogeneity landscape under different dimensionality reduction methods and parameter settings. Columns correspond to the KDE of the **whole sample** (left), followed by **HC**, **MCI**, and **AD** groups (right). **(A)** Results obtained using multidimensional scaling (MDS) with three different random initializations (top three rows), demonstrating consistent landscape organization across seeds. **(B)** Principal coordinates analysis (PCoA), **(C)** principal component analysis (PCA), and **(D)** Isomap embeddings are shown in subsequent rows. Across all methods, the same qualitative structure is preserved: healthy individuals form a dense central concentration, MCI participants show increased dispersion, and AD-diagnosed individuals exhibit the greatest spread and directional heterogeneity. These results confirm that the heterogeneity landscape is not an artifact of a specific embedding method but reflects stable organization in the high-dimensional normative deviation space.

Supplementary References

1. Casey, B. J. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience* **32,** 43–54. 10.1016/j.dcn.2018.03.001 (2018).
2. Di Martino, A. *et al.* The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry* **19,** 659–667. 10.1038/mp.2013.78 (2014).
3. Di Martino, A. *et al.* Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Scientific Data* **4,** 1–15. 10.1038/sdata.2017.10 (2017).
4. Milham, P. M., Damien, F., Maarten, M. & Stewart, H. M. The ADHD-200 Consortium: A model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in Systems Neuroscience* **6,** 1–5. 10.3389/fnsys.2012.00062 (2012).
5. Jack, C. R. *et al.* The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging* **27,** 685–691. 10.1002/jmri.21049 (2008).
6. Petersen, R. C. *et al.* Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. *Neurology* **74,** 201–209. 10.1212/WNL.0b013e3181cb3e25 (2010).
7. ADNI Team. *ADNIMERGE: Alzheimer's Disease Neuroimaging Initiative* 2023.
8. Fowler, C. *et al.* Fifteen Years of the Australian Imaging, Biomarkers and Lifestyle (AIBL) Study: Progress and Observations from 2,359 Older Adults Spanning the Spectrum from Cognitive Normality to Alzheimer's Disease. *Journal of Alzheimer's Disease Reports* **5,** 443–468. 10.3233/ADR-210005 (2021).
9. Ellis, K. A. *et al.* The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *International Psychogeriatrics* **21,** 672–687. 10.1017/S1041610209009405 (2009).
10. Snoek, L. *et al.* The Amsterdam Open MRI Collection, a set of multimodal MRI datasets for individual difference analyses. *Scientific Data* **8,** 85. 10.1038/s41597-021-00870-6 (Mar. 2021).
11. Loughland, C. *et al.* Australian Schizophrenia Research Bank: A database of comprehensive clinical, endophenotypic and genetic data for aetiological studies of schizophrenia. *Australian and New Zealand Journal of Psychiatry* **44,** 1029–1035. 10.3109/00048674.2010.501758 (2010).
12. Holmes, A. J. *et al.* Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Scientific Data* **2,** 1–16. 10.1038/sdata.2015.31 (2015).
13. Buckner, R. L., Roffman, J. L. & Smoller, J. W. *Brain Genomics Superstruct Project (GSP)* 2014. 10.7910/dvn/25833.
14. Salum, G. A. *et al.* High risk cohort study for psychiatric disorders in childhood: rationale, design, methods and preliminary results. *International Journal of Methods in Psychiatric Research* **24,** 58–73. 10.1002/mpr.1459 (Mar. 2015).
15. Shafiei, G. *et al.* Reproducible Brain Charts: An open data resource for mapping brain development and its associations with mental health. *Neuron* **113,** 3758–3779. 10.1016/j.neuron.2025.08.026 (Nov. 2025).
16. Esteban, O., Markiewicz, C. J., Blair, R., Poldrack, R. A. & Gorgolewski, K. J. *sMRIPrep: Structural MRI PREProcessing workflows* Aug. 2023. 10.5281/zenodo.8253909.
17. Halchenko, Y. *et al.* DataLad: distributed system for joint management of code, data, and their relationship. *Journal of Open Source Software* **6,** 3262. 10.21105/joss.03262 (July 2021).

18. Liu, S. *et al.* Chinese Color Nest Project : An accelerated longitudinal brain-mind cohort. *Developmental Cognitive Neuroscience* **52,** 101020. 10.1016/j.dcn.2021.101020 (Dec. 2021).

19. Fan, X. R. *et al.* A longitudinal resource for population neuroscience of school-age children and adolescents in China. *Scientific Data* **10,** 545. 10.1038/s41597-023-02377-8 (Aug. 2023).

20. Zuo, X. N. *et al.* An open science resource for establishing reliability and reproducibility in functional connectomics. *Scientific Data* **1,** 1–13. 10.1038/sdata.2014.49 (2014).

21. Park, D. C. *et al.* The Dallas Lifespan Brain Study: A Comprehensive Adult Lifespan Data Set of Brain and Cognitive Aging. *Scientific Data* **12,** 1–14. 10.1038/s41597-025-04847-7 (2025).

22. Park, D. *et al.* *"The Dallas Lifespan Brain Study"* 2025. doi:10.18112/openneuro.ds004856.v1.3.0.

23. Soh, S. E. *et al.* Cohort profile: Growing up in Singapore towards healthy outcomes (GUSTO) birth cohort study. *International Journal of Epidemiology* **43,** 1401–1409. 10.1093/ije/dyt125 (2014).

24. Alexander, L. M. *et al.* An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific Data* **4,** 170181. 10.1038/sdata.2017.181 (Dec. 2017).

25. Bookheimer, S. Y. *et al.* The Lifespan Human Connectome Project in Aging: An overview. *NeuroImage* **185,** 335–348. 10.1016/j.neuroimage.2018.10.009 (2019).

26. Harms, M. P. *et al.* Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects. *NeuroImage* **183,** 972–984. 10.1016/j.neuroimage.2018.09.060 (2018).

27. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80,** 105–124. 10.1016/j.neuroimage.2013.04.127 (Oct. 2013).

28. Robinson, E. C. *et al.* Multimodal surface matching with higher-order smoothness constraints. *NeuroImage* **167,** 453–465. 10.1016/j.neuroimage.2017.10.037 (2018).

29. Somerville, L. H. *et al.* The Lifespan Human Connectome Project in Development: A large-scale study of brain connectivity development in 5–21 year olds. *NeuroImage* **183,** 456–468. 10.1016/j.neuroimage.2018.08.050 (2018).

30. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: An overview. *NeuroImage* **80,** 62–79. 10.1016/j.neuroimage.2013.05.041 (2013).

31. Robinson, E. C. *et al.* MSM: A new flexible framework for multimodal surface matching. *NeuroImage* **100,** 414–426. 10.1016/j.neuroimage.2014.05.069 (2014).

32. Tobe, R. H. *et al.* A longitudinal resource for studying connectome development and its psychiatric associations during childhood. *Scientific Data* **9,** 300. 10.1038/s41597-022-01329-y (June 2022).

33. LaMontagne, P. J. *et al.* *OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease* Dec. 2019. 10.1101/2019.12.13.19014902.

34. Herrick, R. *et al.* XNAT Central: Open sourcing imaging research data. *NeuroImage* **124,** 1093–1096. 10.1016/j.neuroimage.2015.06.076 (Jan. 2016).

35. Satterthwaite, T. D. *et al.* Neuroimaging of the Philadelphia Neurodevelopmental Cohort. *NeuroImage* **86,** 544–553. 10.1016/j.neuroimage.2013.07.064 (Feb. 2014).

36. Satterthwaite, T. D. *et al.* The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage* **124,** 1115–1119. 10.1016/j.neuroimage.2013.05.012 (Jan. 2016).

37. Wei, D. *et al.* Structural and functional brain scans from the cross-sectional Southwest University adult lifespan dataset. *Scientific Data* **5,** 180134. 10.1038/sdata.2018.134 (July 2018).

38. Hankin, J. H. *et al.* Singapore Chinese Health Study: Development, Validation, and Calibration of the Quantitative Food Frequency Questionnaire. *Nutrition and Cancer* **39,** 187–195. 10.1207/S15327914nc392{_}5 (2001).

39. Qin, S. *et al.* Association between objectively measured, multidimensional sleep health and cognitive function in older adults: cross-sectional wearable tracker study. *Sleep Medicine* **132,** 106569. 10.1016/j.sleep.2025.106569 (2025).

40. Xu, X. *et al.* The SINgapore GERiatric Intervention Study to Reduce Cognitive Decline and Physical Frailty (SINGER): Study Design and Protocol. *The Journal of Prevention of Alzheimer's Disease* **9,** 40–48. 10.14283/jpad.2022.5 (Jan. 2022).

41. Liu, W. *et al.* Longitudinal test-retest neuroimaging data from healthy young adults in southwest China. *Scientific Data* **4,** 170017. 10.1038/sdata.2017.17 (Feb. 2017).

42. Loo, E. X. L. *et al.* Cohort profile: Singapore Preconception Study of Long-Term Maternal and Child Outcomes (S-PRESTO). *European Journal of Epidemiology* **36,** 129–142. 10.1007/s10654-020-00697-2 (Jan. 2021).

43. Chan, S. Y. *et al.* Neonatal Nucleus Accumbens Microstructure Modulates Individual Susceptibility to Preconception Maternal Stress in Relation to Externalizing Behaviors. *Journal of the American Academy of Child & Adolescent Psychiatry* **63,** 1035–1046. 10.1016/j.jaac.2023.12.011 (Oct. 2024).

44. Chopra, S. *et al.* The Transdiagnostic Connectome Project: an open dataset for studying brain-behavior relationships in psychiatry. *Scientific Data* **12,** 1–22. 10.1038/s41597-025-04895-z (2025).

45. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562,** 203–209. 10.1038/s41586-018-0579-z (Oct. 2018).

46. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature Neuroscience* **19,** 1523–1536. 10.1038/nn.4393 (Nov. 2016).

47. 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. 10.1016/j.neuroimage.2017.10.034 (Feb. 2018).

48. Daugherty, A. M. & Raz, N. A virtual water maze revisited: Two-year changes in navigation performance and their neural correlates in healthy adults. *NeuroImage* **146,** 492–506. 10.1016/j.neuroimage.2016.09.044 (Feb. 2017).

49. Rutherford, S. *et al.* The normative modeling framework for computational psychiatry. *Nature Protocols* **17,** 1711–1734. 10.1038/s41596-022-00696-5 (July 2022).

50. Rehak Buckova, B. *et al.* Using normative models pre-trained on cross-sectional data to evaluate intra-individual longitudinal changes in neuroimaging data. *eLife* **13,** 1–33. 10.7554/elife.95823.4 (2025).

51. Hilal, S. *et al.* Markers of Cardiac Dysfunction in Cognitive Impairment and Dementia. *Medicine* **94,** e297. 10.1097/MD.0000000000000297 (Jan. 2015).

52. Hilal, S. *et al.* Cortical cerebral microinfarcts predict cognitive decline in memory clinic patients. *Journal of Cerebral Blood Flow and Metabolism* **40,** 44–53. 10.1177/0271678X19835565 (2020).

53. Chong, J. S. X. *et al.* Influence of cerebrovascular disease on brain networks in prodromal and clinical Alzheimer's disease. *Brain* **140,** 3012–3022. 10.1093/brain/awx224 (Nov. 2017).

54. Fischl, B. *et al.* Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33,** 341–355. 10.1016/S0896-6273(02)00569-X (2002).

55. Fischl, B. FreeSurfer. *NeuroImage* **62,** 774–781. 10.1016/j.neuroimage.2012.01.021 (Aug. 2012).

56. Yeo, B. T. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* **106,** 1125–1165. 10.1152/jn.00338.2011 (2011).

57. Yan, X. *et al.* Homotopic local-global parcellation of the human cerebral cortex from resting-state functional connectivity. *NeuroImage* **273,** 120010. 10.1016/j.neuroimage.2023.120010 (2023).

58. Hartigan, P. M. Algorithm AS 217: Computation of the Dip Statistic to Test for Unimodality. *Applied Statistics* **34,** 320. 10.2307/2347485 (1985).

59. Adolfsson, A., Ackerman, M. & Brownstein, N. C. To cluster, or not to cluster: An analysis of clusterability methods. *Pattern Recognition* **88,** 13–26. 10.1016/j.patcog.2018.10.026 (2019).