nature communications

Article | Article in Press

https://doi.org/10.1038/s41467-026-79269-z

Mapping hippocampal subfield heterogeneity across neurological disorders by normative modeling

Received: 22 November 2025

Accepted: 14 October 2026

Published online: 03 November 2026

Check for updates

A list of authors and their affiliations appears at the end of the paper

The hippocampus serves as a hub of cognition, and its subfields differ in their susceptibility to aging and neurological disorders. Here, we established normative references for hippocampal and subfield volumes using 3D T1-weighted magnetic resonance imaging (MRI) scans from 34,789 healthy individuals (range from 4 to 85 years), revealing varied peak ages (27.4-40.1 years) across subfields. We then mapped the hippocampus and its subfields across various neurological disorders (mild cognitive impairment, Alzheimer's disease, Parkinson's disease, cerebral small vessel disease, and multiple sclerosis) by deriving personalized deviation (centile) scores from these references in 3,521 patients. These centile scores captured disease-specific hippocampal abnormalities, supported disease classification, and were associated with cognitive performance, longitudinal cognitive outcomes, and disease progression across disorders. Finally, a personalized hippocampal health report was developed to facilitate clinical translation. These findings demonstrate the value of hippocampal and subfield normative references for personalized diagnosis, cognitive assessment, and prognostic evaluation across neurological diseases.

The hippocampus, which is located in the medial temporal lobe, plays a key role in learning and memory functions[1]. It undergoes structural changes during development and aging and is vulnerable to physiological and pathological stressors[2,3]. Hippocampal atrophy has been observed in a range of neurological disorders associated with cognitive impairment, including mild cognitive impairment (MCI)[4], Alzheimer's disease (AD)[5], Parkinson's disease (PD)[6], cerebral small vessel disease (CSVD)[7], and multiple sclerosis (MS)[8]. Because hippocampal volume loss may precede overt clinical symptoms, it has been proposed as a potential biomarker for early diagnosis and monitoring cognitive decline[9]. The hippocampus comprises distinct subfields that differ in cytoarchitecture, connectivity, and function, showing heterogeneous vulnerability across normal aging and diseases[10]. Recent studies have focused on hippocampal subfields to reveal disease-specific atrophy patterns and to improve understanding of structural heterogeneity in neurological disorders[11,12].

High-resolution MRI combined with automated segmentation enables reliable quantification of hippocampal subfield volumes in large-scale datasets[13]. Hippocampal subfield atrophy has been consistently reported in neurological disorders and is associated with cognitive performance. Atrophy of the subiculum and presubiculum in MCI and AD is closely linked to memory decline[14], whereas degeneration of the Cornu Ammonis (CA) subfields has been associated with episodic memory and language impairment in PD, CSVD, and MS[7,15,16]. These findings highlight the heterogeneity in hippocampal subfields and suggest distinct pathological mechanisms underlying cognitive dysfunction[17,18]. Nevertheless, existing studies are limited by case-control designs, inconsistent methodological approaches, and insufficient consideration of age-related trajectories, which constrain reproducibility and clinical applicability[19–25].

Recent years have seen a rapid expansion of normative modeling approaches and the development of large-scale lifespan brain charts that characterize structural trajectories across the lifespan[26,27]. These

e-mail: demaodeng@163.com; rongjunyu@hkbu.edu.hk; ndyfy02301@ncu.edu.cn; 18612161210@163.com; xinian.zuo@bnu.edu.cn; liuyaou@bjtth.org

https://doi.org/10.1038/s41467-026-79269-z

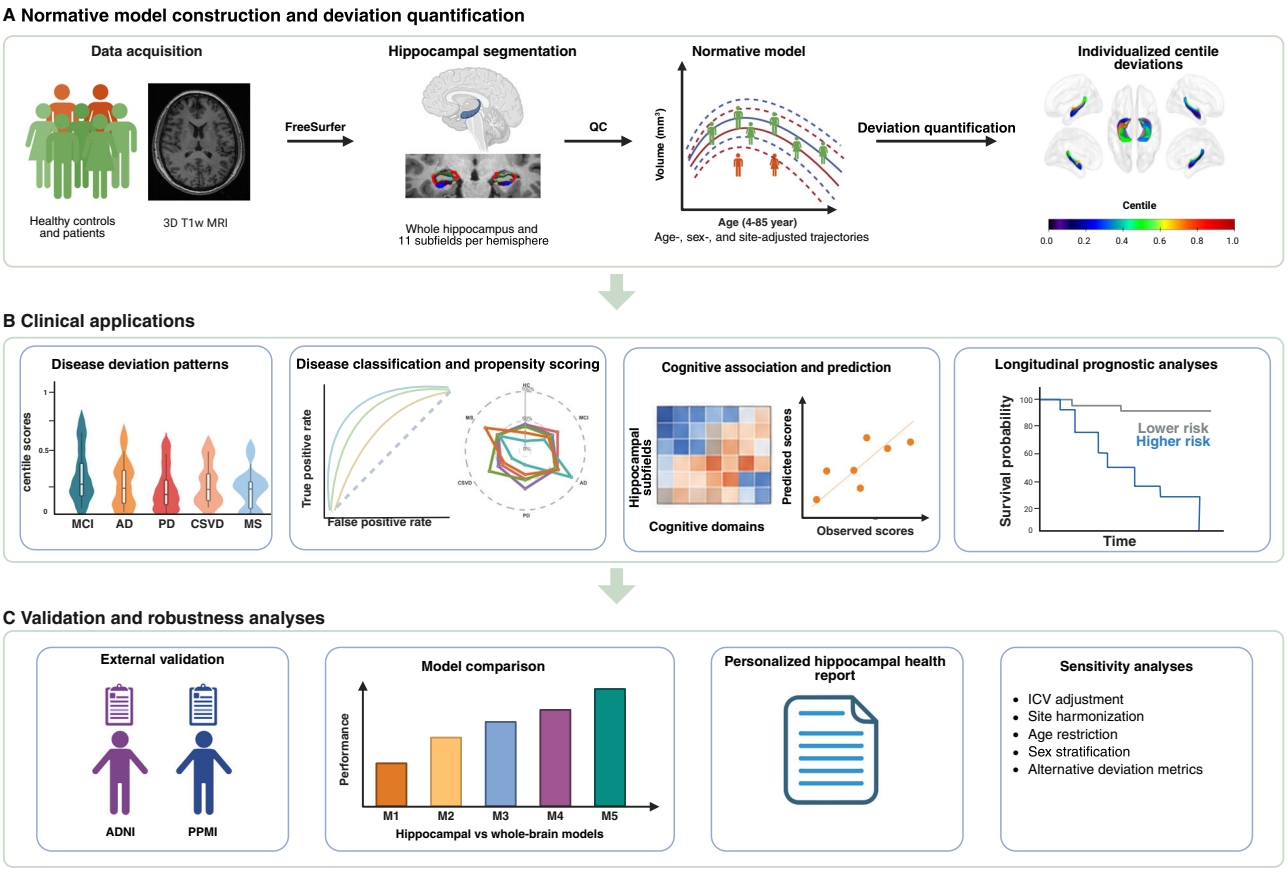


Fig. 1 | Overview of the study design and analytical framework. A Multicenter 3D T1w-MRI data from healthy controls and patient cohorts were processed using FreeSurfer-based hippocampal and subfield segmentation, followed by quality control. Lifespan normative models adjusted for age, sex, and acquisition site were then constructed to derive individualized centile deviation scores. **B** The individualized centile scores were used to characterize disease-specific hippocampal deviation patterns and to support disease classification, propensity scoring, cognitive association and prediction analyses, and longitudinal prognostic modeling. **C** Validation and robustness analyses included external validation in independent cohorts, model comparison, individualized deviation profiling, and sensitivity analyses. Created in BioRender. DU, X. (2026) https://BioRender.com/fygi5m6. T1w T1-weighted, QC quality control, ADNI Alzheimer's Disease Neuroimaging Initiative, PPMI Parkinson's Progression Markers Initiative, ICV intracranial volume.

models generate deviation scores that separate disease-related alterations from normal aging and provide individualized neuroanatomical assessments relative to population references, thereby supporting improved diagnostic precision and personalized monitoring[28,29]. While such approaches have been applied to global or regional brain measures, including whole hippocampal volume, subfield-level lifespan references remain to be systematically established. Moreover, it remains unclear whether subfield-level characterization provides added value beyond broader brain measures for clinical applications[30].

In this study, we aimed to develop lifespan normative references for hippocampal and subfield volumes and to evaluate their utility across neurological disorders, including MCI, AD, PD, CSVD, and MS (Fig. 1). Using a large-scale, multicenter 3D T1w-MRI dataset comprising 34,789 healthy controls (HCs) from large-scale international cohorts, we constructed normative references across ages 4–85 years and derived individualized centile scores. We then applied these scores to characterize disease-related hippocampal deviation patterns and evaluate their associations with cognitive performance, disease classification, and longitudinal outcomes. To assess generalizability and added value, we performed independent external validation and benchmarked hippocampal models against whole-hippocampus and whole-brain models across downstream tasks. Finally, we developed an illustrative personalized hippocampal health report to summarize individual deviation profiles and support individualized interpretation.

Results

Normative references of hippocampal and subfield volumes

We established normative references for bilateral whole-hippocampal and hippocampal subfield volumes using 34,789 HCs aged 4–85 years from multiple international cohorts (Table 1 and Table S1). Hippocampal segmentation was performed using FreeSurfer (version 7.3.2), including the whole hippocampus and 11 anatomical subfields in each hemisphere (CA1, CA3, CA4, granule cell layer of the dentate gyrus [GC-DG], molecular layer [ML], presubiculum, subiculum, parasubiculum, hippocampal-amygdaloid transition area [HATA], fimbria, and tail). Normative references were fitted in HCs using generalized additive models for location, scale, and shape (GAMLSS), considering age and sex as fixed effects, and site as a random effect[31,32].

The hippocampal volume exhibited an inverted U-shaped trajectory across ages 4–85 years, increasing rapidly during development, peaking at 33–34 years (left, 33.3 years [95% confidence interval (CI) 32.5–34.1]; right, 34.0 years [95% CI 33.7–34.4]), and decreasing gradually thereafter. Most hippocampal subfields showed broadly similar inverted U-shaped trajectories, although the timing of peak volume varied across anatomical regions (Fig. 2A). Across bilateral subfields, estimated peak ages ranged from 27.4 to 40.1 years, excluding the parasubiculum, whose fitted trajectory showed an approximately monotonic decrease from childhood onward and therefore reached its maximum at the lower boundary of the modeled age range. The earliest post-childhood peak was observed in the left fimbria at 27.4 years,

https://doi.org/10.1038/s41467-026-79269-z

Table 1 | Demographic and clinical characteristics of participants

	HC (*n* = 34,789)	MCI (*n* = 308)	AD (*n* = 604)	PD (*n* = 1226)	CSVD (*n* = 864)	MS (*n* = 519)
Age, years (mean [SD])	37.2 (19.33)	64.85 (8.25)	68.15 (8.23)	61.26 (9.45)	60.57 (10.08)	36.28 (12.07)
Female, n (%)	17,866 (51.4)	193 (62.66)	360 (59.60)	558 (45.51)	363 (42.01)	338 (65.13)
Education, years (median [IQR])	15.00 [12.00, 16.00]	11.00 [9.00, 13.00]	9.00 [8.00, 12.00]	12.00 [9.00, 16.00]	9.00 [6.00, 12.00]	15.00 [11.00, 16.00]
MMSE (median [IQR])	29.00 [27.00, 30.00]	26.00 [24.00, 28.00]	18.00 [12.00, 23.00]	27.00 [24.00, 28.00]	26.00 [23.00, 28.00]	NA
MoCA (median [IQR])	26.00 [24.00, 28.00]	21.00 [19.00, 23.00]	13.00 [7.00, 17.00]	22.00 [18.25, 25.00]	21.00 [17.00, 25.00]	NA
BVMT-R (median [IQR])	NA	NA	NA	NA	NA	35.00 [26.00, 48.00]
SDMT (median [IQR])	NA	NA	NA	NA	NA	48.50 [36.00, 58.25]
PASAT (median [IQR])	NA	NA	NA	NA	NA	45.00 [34.25, 50.00]
EDSS (median [IQR])	NA	NA	NA	NA	NA	2.00 [1.00, 3.50]

n number, *SD* standard deviation, *IQR* interquartile range, *HC* healthy control, *MCI* mild cognitive impairment, *AD* Alzheimer's disease, *PD* Parkinson's disease, *CSVD* cerebral small vessel disease, *MS* multiple sclerosis, *MMSE* Mini-Mental State Examination, *MoCA* Montreal Cognitive Assessment, *BVMT-R* Brief Visuospatial Memory Test-Revised, *PASAT* Paced Auditory Serial Addition Test, *SDMT* Symbol Digit Modalities Test, *EDSS* Expanded Disability Status Scale.

whereas the latest peak was observed in the right CA3 at 40.1 years. Several subfields showed modest hemispheric differences in peak timing (Fig. 2B).

To further characterize age-related hippocampal and subfield volume changes, we first computed the first derivative of the fitted median trajectories to estimate absolute rates of change (mm^3/year). We then derived relative rates of change (%/year) by normalizing the absolute rates to the predicted median volume at each age, enabling comparison across subfields of different sizes. These cross-sectionally inferred rates revealed progressively slowing growth approaching the peak age, followed by accelerating decline thereafter (Fig. 2C, D).

Hippocampal and subfield deviations across neurological disorders

Personalized centile scores were derived by comparing each patient's hippocampal and hippocampal subfield volumes with age-, sex-, and site-adjusted normative references. Case-control differences in centile scores were assessed between each disease group and age-, sex-, and site-matched HCs using Monte Carlo permutation tests based on Welch's t statistic, with two-sided permutation P values estimated using 5000 permutations. Cohen's d with 95% confidence intervals was calculated to quantify effect sizes. Multiple comparisons were controlled within each disease using the Benjamini–Hochberg false discovery rate (FDR) procedure. To highlight the most sensitive disease-related deviations, the three subfields with the largest effect sizes in each group were reported. Detailed group comparisons are presented in Fig. 3 and Table S4.

All disease groups showed distinct patterns of hippocampal and subfield deviation. For bilateral whole-hippocampal centile scores, MCI, AD, PD, CSVD, and MS all showed reductions relative to matched HCs, with Cohen's d values ranging from 0.08 to 0.94 and all comparisons surviving FDR correction. At the subfield level, patients with MCI showed significant reductions in 15 of 22 subfields, with the largest effects observed in the right parasubiculum (Cohen's $d = 0.34$), left presubiculum (Cohen's $d = 0.30$), and left CA4 (Cohen's $d = 0.27$; all FDR-adjusted $P < 0.001$). AD showed widespread reductions across all 22 subfields, with the strongest effects in the right subiculum (Cohen's $d = 1.00$), left subiculum (Cohen's $d = 0.98$), and right molecular layer (Cohen's $d = 0.93$; all FDR-adjusted $P < 0.001$). PD showed significant reductions in 18 of 22 subfields, most prominently in the left presubiculum (Cohen's $d = 0.28$), right presubiculum (Cohen's $d = 0.27$), and left HATA (Cohen's $d = 0.25$; all FDR-adjusted $P < 0.001$). CSVD showed significant reductions in 13 of 22 subfields, with the largest effects in the right hippocampal tail (Cohen's $d = 0.29$), left hippocampal tail (Cohen's $d = 0.24$), and left presubiculum (Cohen's $d = 0.24$; all FDR-adjusted $P < 0.001$). MS showed significant reductions across all 22 subfields, with the strongest effects in the right hippocampal tail (Cohen's $d = 0.68$), left hippocampal tail (Cohen's $d = 0.68$), and left HATA (Cohen's $d = 0.53$; all FDR-adjusted $P < 0.001$).

The centile Mahalanobis distance (CMD), a summary measure of multivariate hippocampal abnormality, was significantly increased in AD (Cohen's $d = 0.35$), CSVD (Cohen's $d = 0.16$), and MS (Cohen's $d = 0.69$; all FDR-adjusted $P < 0.01$), whereas no significant CMD differences were observed in MCI or PD after FDR correction.

Disease classification and disease propensity score estimation by centile scores

Using bilateral whole-hippocampal and hippocampal subfield centile scores as the primary hippocampal feature set, we trained disease-specific binary classifiers to distinguish each disease group from age- and sex-matched HCs. Candidate classifiers included regularized logistic regression and single-hidden-layer neural network models, with final models selected within the training set and evaluated in an independent held-out test set. Comparisons with alternative

https://doi.org/10.1038/s41467-026-79269-z

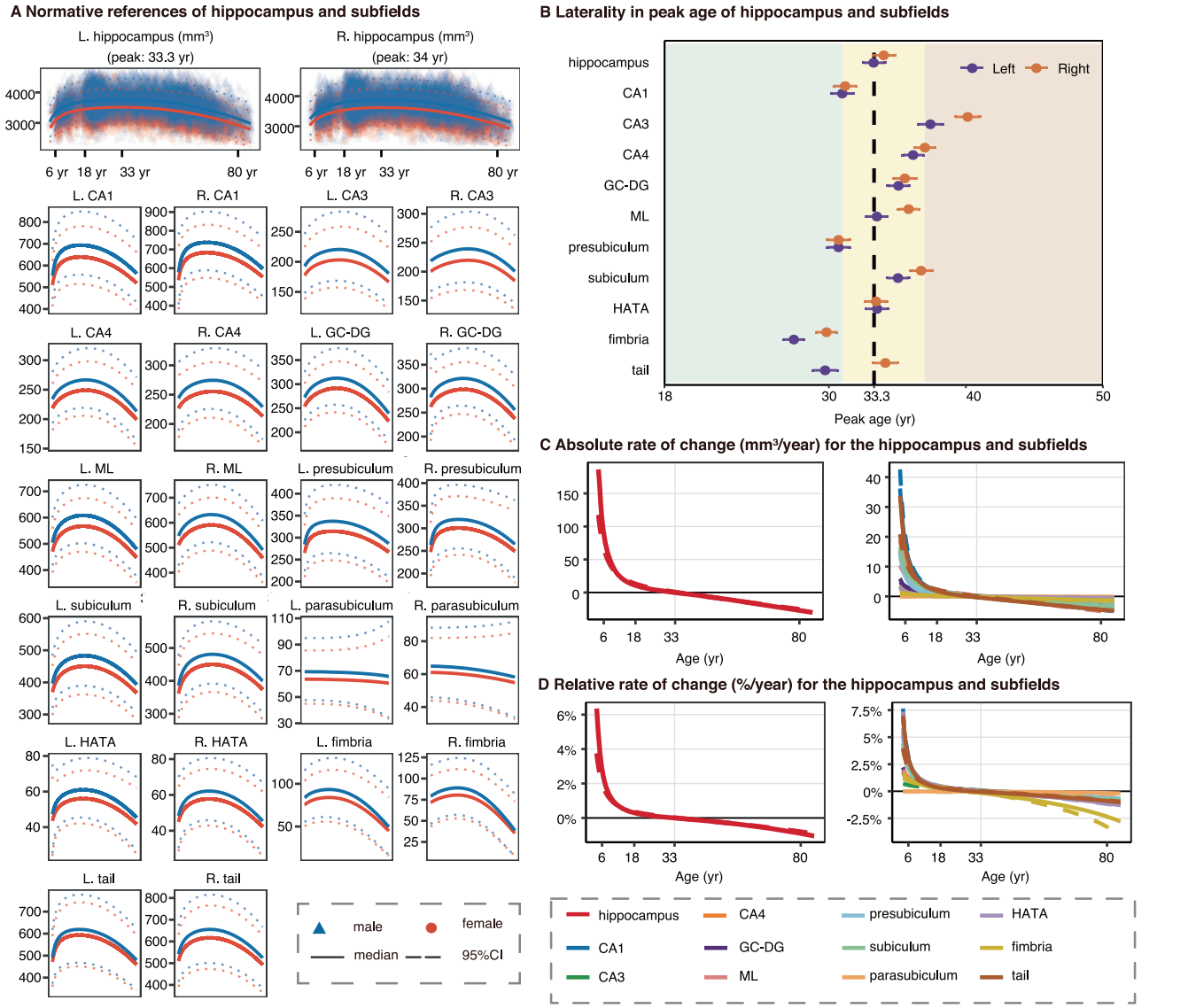


Fig. 2 | Normative lifespan trajectories of hippocampal subfield volumes. Analyses were based on 34,789 healthy individuals. **A** Fitted median trajectories for the whole hippocampus and 11 hippocampal subfields in each hemisphere across the lifespan, with dotted lines indicating the corresponding 95% reference intervals. **B** Hemispheric differences in estimated peak age for the whole hippocampus and hippocampal subfields. Points indicate the estimated peak ages, and horizontal bars indicate the corresponding 95% confidence intervals. The vertical dashed line marks the peak age of the whole hippocampus as a reference. **C, D** Absolute rates of change (mm³/year) and relative rates of change (%/year), respectively, across age, derived from the first derivative of the fitted median normative trajectories. In (**C**, **D**), solid and dashed lines denote the left and right hemispheres, respectively. CA Cornu Ammonis, GC-DG granule cell–dentate gyrus, ML molecular layer, HATA hippocampal-amygdaloid transition area, RI reference interval, CI confidence interval. Source data are provided as a Source Data file.

hippocampal and whole-brain feature sets are reported in the subsequent benchmark analysis and detailed in Table S2.

Disease-specific classification models achieved area under the receiver operating characteristic curve (AUC) values of 0.89 [95% CI 0.85–0.94] for MCI, 0.86 [95% CI 0.82–0.89] for AD, 0.74 [95% CI 0.71–0.78] for PD, 0.80 [95% CI 0.76–0.84] for CSVD, and 0.91 [95% CI 0.88–0.94] for MS (Fig. 4A). The corresponding held-out test-set confusion matrices are shown in Fig. 4B.

The estimated disease propensity score (DPS) values showed the expected disease-specific pattern, with each disease model yielding the highest DPS in its corresponding diagnostic group (Fig. 4C). Median DPS values (interquartile range [IQR]) were 0.76 [0.55–0.87] for MCI, 0.85 [0.63–0.92] for AD, 0.68 [0.53–0.80] for PD, 0.70 [0.53–0.82] for CSVD, and 0.93 [0.78–0.99] for MS. In additional pairwise disease classification analyses, held-out test-set AUCs ranged from 0.68 to 0.99 across disease pairs (Fig. 4D).

Associations and predictive utility of hippocampal and subfield centile scores for cognitive performance in cross-sectional data

To examine the associations between hippocampal and subfield centile scores and cognitive performance, we performed partial Spearman correlation analyses adjusted for age and sex. Predictive utility was further evaluated using support vector regression (SVR) with a radial basis function kernel under 10-fold cross-validation, including age, sex, and imaging-derived centile features as predictors. Prediction performance was assessed by Pearson correlations between observed and predicted cognitive scores across pooled held-out fold predictions. Feature-set comparisons for cognitive prediction are reported in the subsequent benchmark analysis and detailed in Table S2.

Higher centile scores were associated with better cognitive performance across disorders (Fig. 5A). In patients with MCI, AD, PD, and CSVD, centile scores showed significant positive correlations with global cognition as measured by the Montreal Cognitive Assessment

https://doi.org/10.1038/s41467-026-79269-z

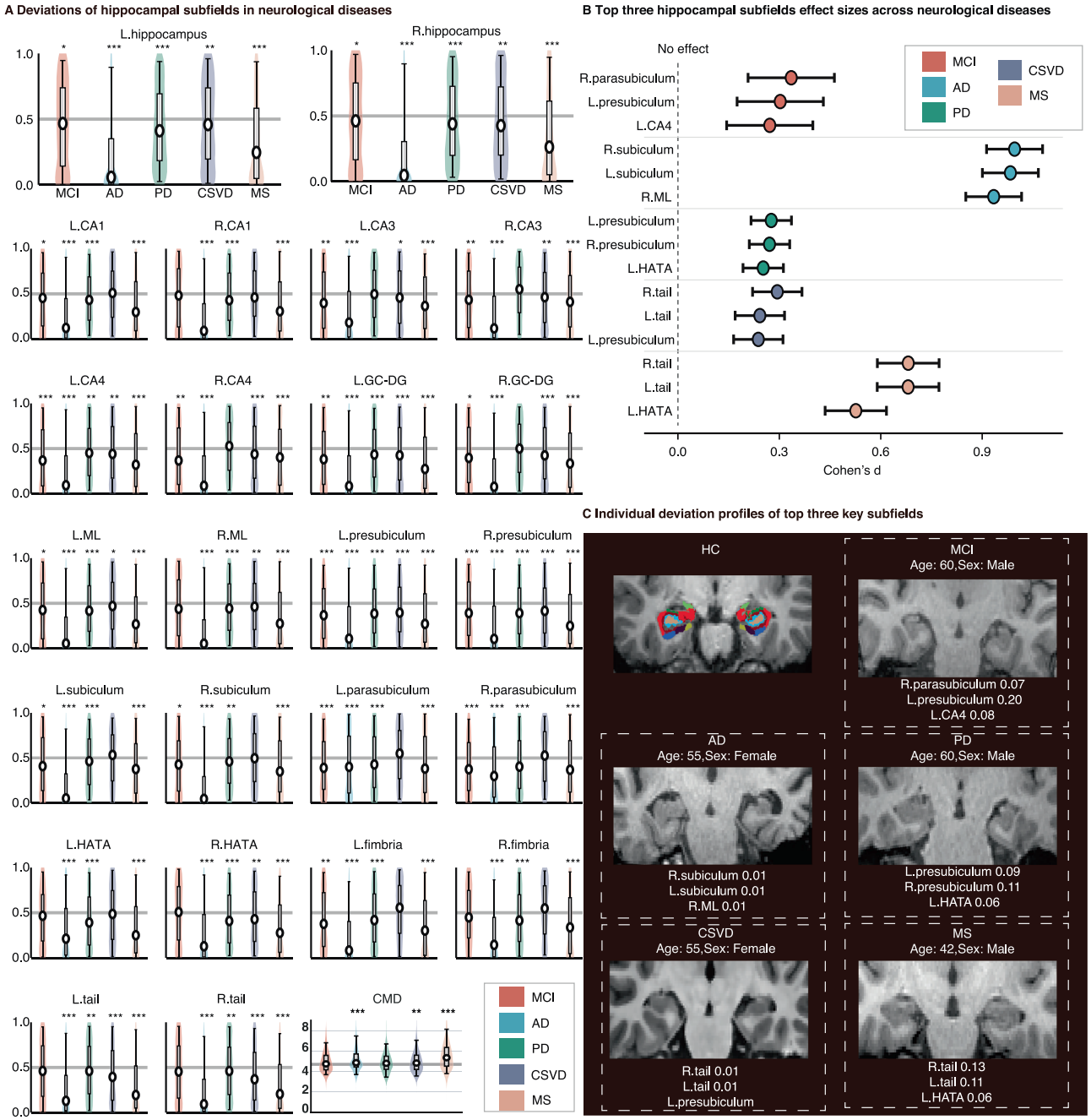


Fig. 3 | Hippocampal subfield heterogeneity across neurological disorders. **A** Distributions of individualized centile scores for the bilateral whole hippocampus and 11 bilateral hippocampal subfields across five neurological disorders. Violin plots show kernel density distributions; open circles indicate medians, and box-plots show the 25th–75th percentiles with whiskers extending from the 5th to the 95th percentiles. The gray horizontal line at 0.5 indicates the normative median. Group differences were assessed using two-sided 1:1 matched Monte Carlo permutation tests based on Welch's t statistic. Sample sizes were MCI, $n = 308$; AD, $n = 604$; PD, $n = 1226$; CSVD, $n = 864$; and MS, $n = 519$. P values were adjusted across 24 hippocampal ROIs within each disease group using the Benjamini–Hochberg FDR procedure. Asterisks indicate FDR-adjusted P value thresholds (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$); exact P values are provided in the Source Data file. CMD is shown as a complementary multivariate summary measure. **B** The three hippocampal ROIs with the largest Cohen's d effect sizes for each neurological disorder. Points indicate Cohen's d estimates, and horizontal lines indicate the corresponding 95% confidence intervals. **C** Representative individual deviation profiles for patients with MCI, AD, PD, CSVD, and MS. Numbers indicate centile scores for the highlighted ROIs. HC healthy control, MCI mild cognitive impairment, AD Alzheimer's disease, PD Parkinson's disease, CSVD cerebral small vessel disease, MS multiple sclerosis, CMD centile Mahalanobis distance, ROI region of interest, FDR false discovery rate, CI confidence interval. Source data are provided as a Source Data file.

(MoCA; $\rho = 0.08–0.43$) and the Mini-Mental State Examination (MMSE; $\rho = 0.08–0.43$; all FDR-adjusted $P < 0.05$). Notably, the strongest associations for each disease were localized to distinct subfields. For the MoCA, peak subfield correlations were observed in the left presubiculum in MCI ($\rho = 0.24$), the right fimbria in AD ($\rho = 0.43$), the right molecular layer in PD ($\rho = 0.20$), and the left subiculum in CSVD ($\rho = 0.26$). For the MMSE, the strongest subfield correlations were localized to the right presubiculum in MCI ($\rho = 0.27$), the right fimbria in AD ($\rho = 0.43$), the left GC-DG in PD ($\rho = 0.16$), and the left GC-DG in CSVD ($\rho = 0.23$). In patients with MS, subfield centile scores were significantly correlated with information processing speed on the Symbol Digit Modalities Test (SDMT; $\rho = 0.20–0.49$), visuospatial memory on

https://doi.org/10.1038/s41467-026-79269-z

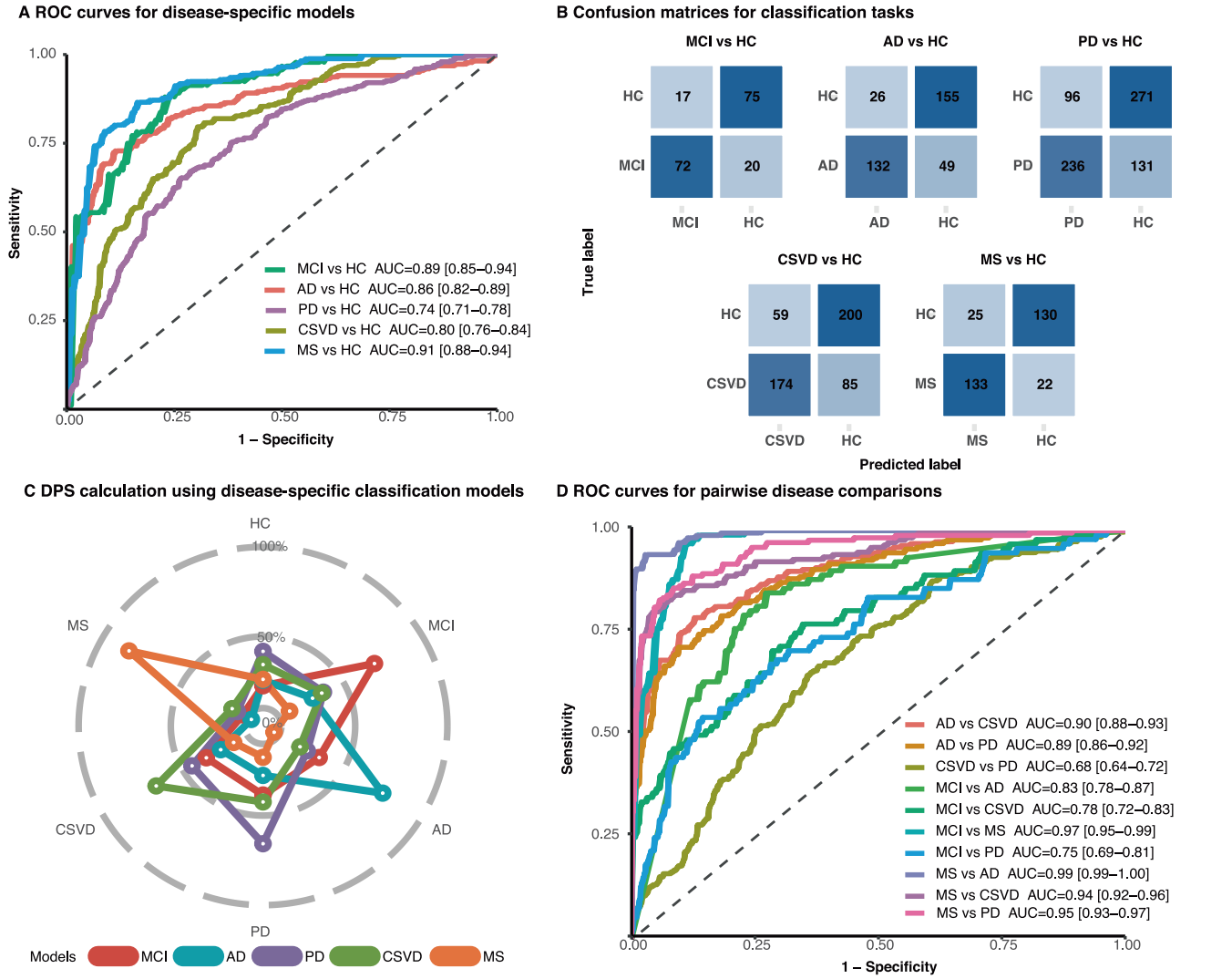


Fig. 4 | Classification performance and disease propensity score estimation. **A** ROC curves for disease-specific one-vs-HC classification models, with AUC and 95% confidence intervals shown for each model. **B** Confusion matrices showing classification performance across the five disease-specific one-vs-HC tasks. Rows indicate true labels, and columns indicate predicted labels. **C** DPS profiles estimated from MCI-, AD-, PD-, CSVD-, and MS-specific classifiers, showing the model-derived propensity of each diagnostic group across disease categories. **D** Pairwise ROC analyses evaluating classification performance between neurological disease groups. ROC receiver operating characteristic, AUC area under the curve, CI confidence interval, DPS disease propensity score, HC healthy control, MCI mild cognitive impairment, AD Alzheimer's disease, PD Parkinson's disease, CSVD cerebral small vessel disease, MS multiple sclerosis.

the Brief Visuospatial Memory Test-Revised (BVMT-R; $\rho = 0.20–0.28$), and working memory on the Paced Auditory Serial Addition Test (PASAT; $\rho = 0.17–0.21$; all FDR-adjusted $P < 0.05$). The strongest MS associations were observed in the left fimbria for SDMT ($\rho = 0.49$), the right CA4 for BVMT-R ($\rho = 0.28$), and the left subiculum for PASAT ($\rho = 0.21$).

The prediction models demonstrated consistent performance across cognitive outcomes (Fig. 5B). Across the MCI, AD, PD, and CSVD groups, correlations between predicted and observed global cognitive scores ranged from 0.43 to 0.64 for the MoCA and from 0.35 to 0.64 for the MMSE, with all associations surviving FDR correction. In the MS group, prediction performance was strongest for PASAT ($r = 0.70$), followed by SDMT ($r = 0.65$) and BVMT-R ($r = 0.56$). These associations were also significant after FDR correction.

Prediction of cognitive decline and disease progression by centile scores in longitudinal data

We further applied the SVR framework to predict follow-up cognitive scores using hippocampal and subfield centile scores (Fig. 5C). Across MCI, AD, and CSVD, correlations between predicted and observed follow-up scores ranged from 0.270 to 0.780 for the MoCA and from 0.350 to 0.600 for the MMSE, with all associations remaining significant after FDR correction. We next evaluated the prognostic utility of hippocampal subfield centile scores using Cox proportional hazards models and Kaplan–Meier risk stratification analyses (Fig. 5D). Kaplan–Meier analyses revealed significantly higher event rates in the high-risk group for HC-to-MCI conversion (log-rank $P = 0.004$), MCI-to-AD conversion (log-rank $P = 0.001$), Expanded Disability Status Scale (EDSS) worsening in MS (log-rank $P = 0.031$), and secondary progressive multiple sclerosis (SPMS) conversion in MS (log-rank $P = 0.004$). Feature-set comparisons for longitudinal cognitive prediction and disease progression analyses are reported in the subsequent benchmark analysis and detailed in Table S2.

Independent external validation

To assess out-of-sample generalizability, we performed external validation using two independent cohorts: the Alzheimer's Disease Neuroimaging Initiative (ADNI) for the Alzheimer's disease continuum and

https://doi.org/10.1038/s41467-026-79269-z

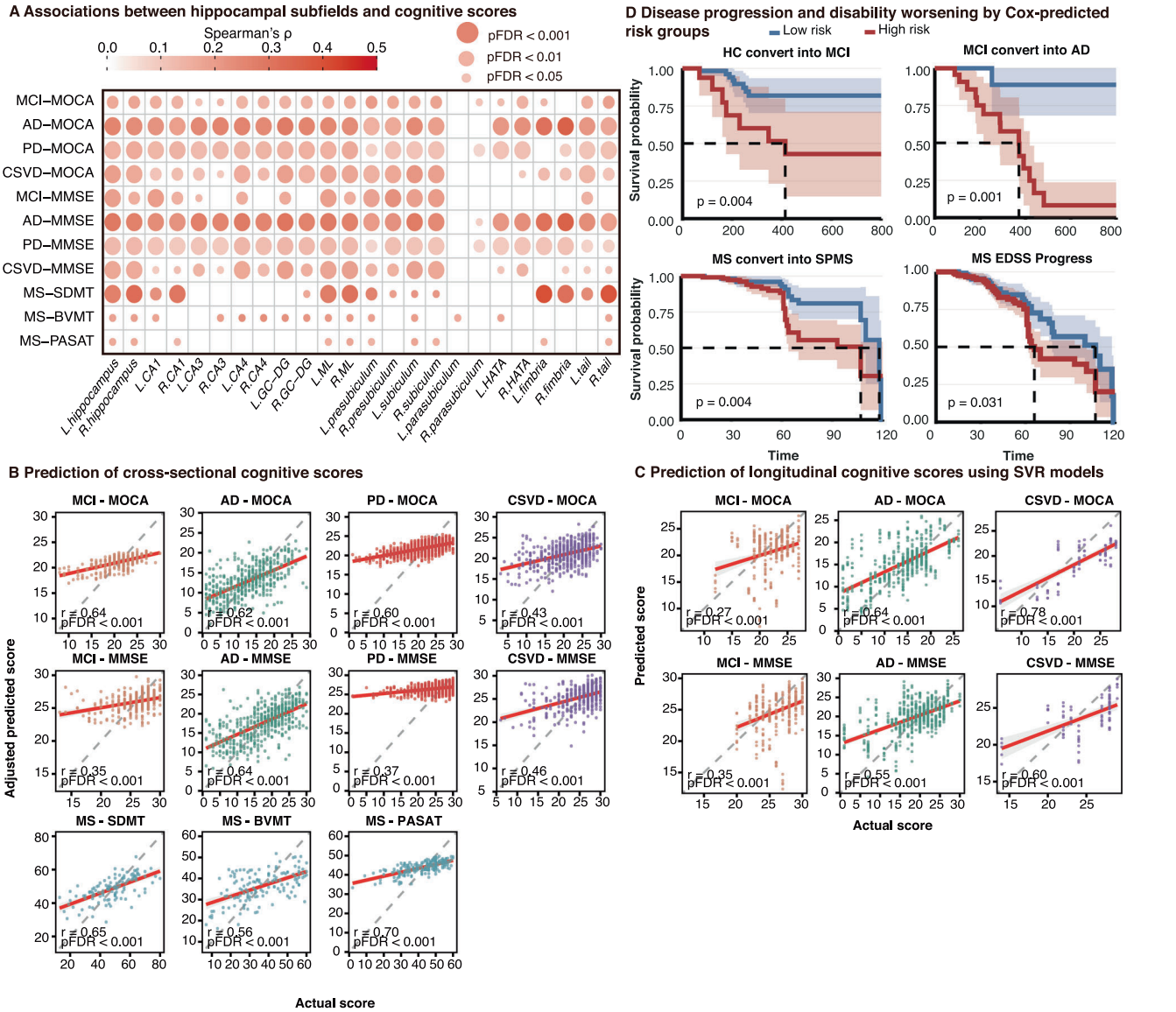


Fig. 5 | Cognitive associations and clinical predictive utility of hippocampal subfield centile scores. A Associations between hippocampal subfield centile scores and cognitive measures across neurological disorders, shown as partial Spearman correlation coefficients (ρ) controlling for age and sex. Color intensity indicates the magnitude of ρ, and circle size indicates FDR-adjusted significance levels (P_FDR < 0.05, P_FDR < 0.01, and P_FDR < 0.001). All tests were two-sided with Benjamini–Hochberg FDR correction. **B, C** Cross-sectional and longitudinal prediction of cognitive outcomes using support vector regression. Scatterplots show predicted versus observed cognitive scores, with Pearson correlation coefficients (r) and two-sided FDR-adjusted P values. Red lines indicate fitted linear regression lines, shaded bands indicate 95% confidence intervals, and gray dashed lines indicate the identity line (y = x). **D** Kaplan–Meier curves for disease progression and disability worsening stratified by the median out-of-fold Cox-predicted risk score. Outcomes included HC-to-MCI conversion, MCI-to-AD conversion, SPMS conversion, and EDSS worsening. Survival distributions were compared using two-sided log-rank tests; shaded bands indicate 95% confidence intervals. Exact P values and sample sizes for all analyses are provided in the Source Data file. ρ Spearman correlation coefficient, r Pearson correlation coefficient, SVR support vector regression, FDR false discovery rate, MoCA Montreal Cognitive Assessment, MMSE Mini-Mental State Examination, SDMT Symbol Digit Modalities Test, BVMT-R Brief Visuospatial Memory Test-Revised, PASAT Paced Auditory Serial Addition Test, HC healthy control, MCI mild cognitive impairment, AD Alzheimer's disease, PD Parkinson's disease, CSVD cerebral small vessel disease, MS multiple sclerosis, SPMS secondary progressive multiple sclerosis, EDSS Expanded Disability Status Scale. Source data are provided as a Source Data file.

the Parkinson's Progression Markers Initiative (PPMI) for Parkinson's disease. Baseline characteristics of the external validation cohorts are provided in Table S5.

In ADNI, new-site recalibrated HCs showed hippocampal region of interest (ROI) centile profiles close to the expected normative distribution, whereas patients with MCI and AD showed progressively lower centile scores across hippocampal regions (Fig. 6A). This graded pattern from HC to MCI to AD was consistent with increasing hippocampal involvement along the Alzheimer's disease continuum. To further examine heterogeneity within the MCI spectrum, we separately analyzed early mild cognitive impairment (EMCI) and late mild cognitive impairment (LMCI) in ADNI. LMCI showed lower hippocampal and subfield centile scores than EMCI, and pairwise classification among HC, EMCI, and LMCI showed discriminative performance, supporting stage-related heterogeneity of hippocampal involvement within MCI (Figure S7). Using the same classification framework as in the internal analyses, external discrimination in ADNI achieved AUCs of 0.781 (95% CI, 0.741–0.821) for MCI versus HC and 0.910 (95% CI, 0.883–0.938) for AD versus HC (Fig. 6B). Hippocampal centile scores also showed external utility for cognitive and prognostic outcomes.

https://doi.org/10.1038/s41467-026-79269-z

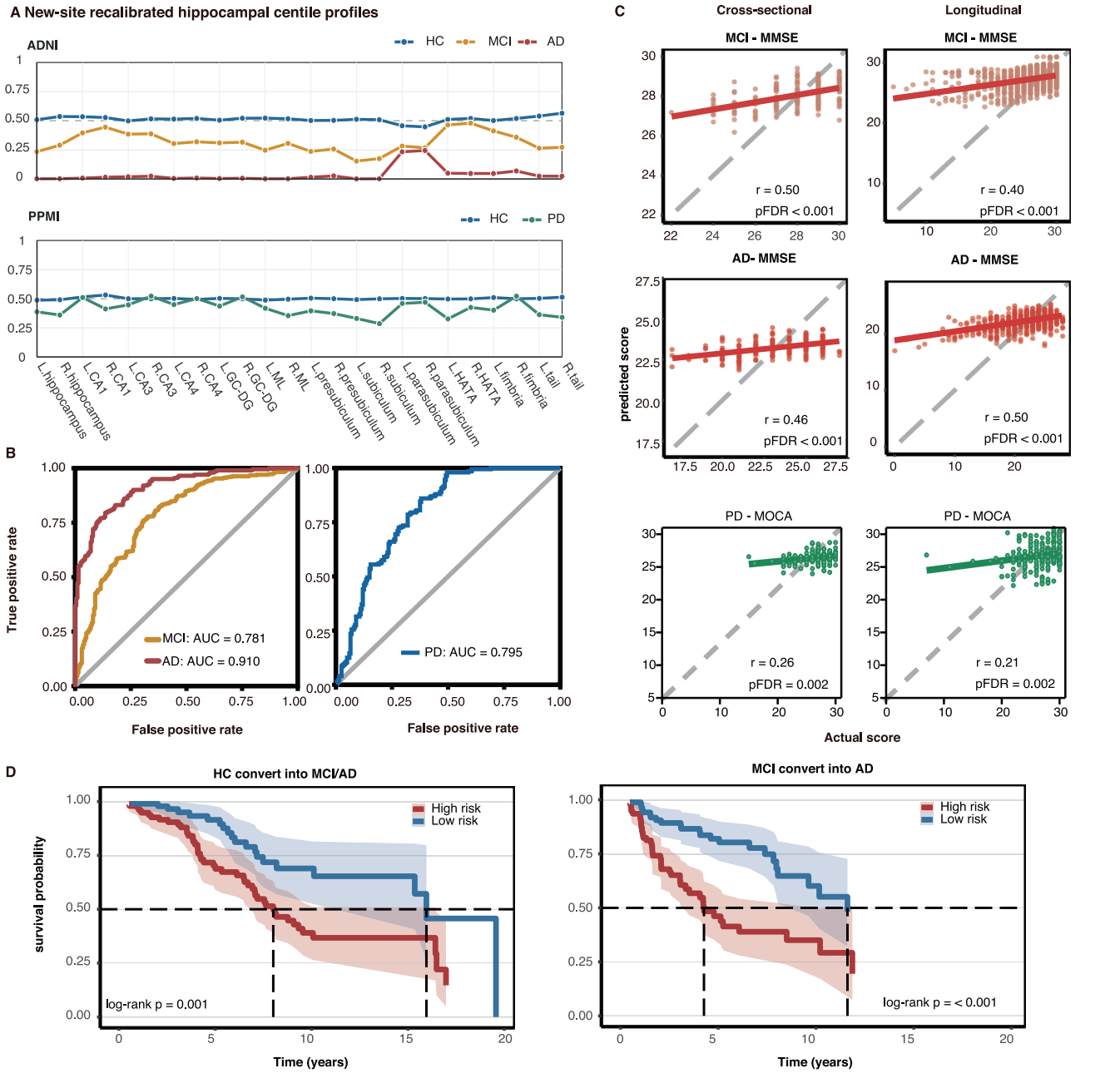


Fig. 6 | External validation of hippocampal centile scores in ADNI and PPMI. **A** New-site-recalibrated hippocampal ROI centile profiles in ADNI and PPMI. Lines indicate median centile scores across hippocampal regions, and the gray horizontal line marks the expected 50th centile in healthy controls. **B** External disease classification performance in ADNI and PPMI. ROC curves show discrimination of MCI versus HC and AD versus HC in ADNI, and PD versus HC in PPMI. AUCs and corresponding 95% confidence intervals are shown in each panel. **C** External cognitive prediction performance. Scatterplots show predicted versus observed MMSE scores for MCI and AD in ADNI and MoCA scores for PD in PPMI, separately for cross-sectional and longitudinal analyses. Red lines indicate fitted linear regression lines, shaded bands indicate 95% confidence intervals, and gray dashed lines indicate the identity line (y = x). Pearson correlation coefficients (*r*) and corresponding two-sided *P* values are shown in each panel. **D** External prognostic validation in ADNI. Kaplan–Meier curves show Cox model-based risk stratification for HC-to-MCI/AD and MCI-to-AD conversion. Survival distributions were compared using two-sided log-rank tests; shaded bands indicate 95% confidence intervals. Exact *P* values and sample sizes for the external classification, cognitive prediction, and prognostic analyses are provided in the Source Data file. AD Alzheimer's disease, ADNI Alzheimer's Disease Neuroimaging Initiative, AUC area under the curve, CI confidence interval, HC healthy control, MCI mild cognitive impairment, MMSE Mini-Mental State Examination, MoCA Montreal Cognitive Assessment, PD Parkinson's disease, PPMI Parkinson's Progression Markers Initiative, ROC receiver operating characteristic. Source data are provided as a Source Data file.

Cross-sectional MMSE prediction achieved correlations of $r = 0.50$ in MCI and $r = 0.46$ in AD, whereas longitudinal MMSE prediction achieved correlations of $r = 0.40$ in MCI and $r = 0.50$ in AD (all FDR-adjusted $P < 0.001$; Fig. 6C). Prognostic analyses further showed significant risk stratification for conversion from HC to MCI or AD (log-rank $P = 0.001$) and from MCI to AD (log-rank $P = 2.09 \times 10^{-5}$; Fig. 6D).

In PPMI, new-site recalibrated HCs showed approximately centered centile distributions, whereas patients with PD showed lower hippocampal centile profiles than external HCs across multiple hippocampal regions (Fig. 6A). These ROI-level patterns were broadly consistent with the internal PD cohort, supporting partial replication of PD-related hippocampal deviations. External discrimination of PD

from HCs achieved an AUC of 0.795 (95% CI, 0.750–0.838; Fig. 6B). External MoCA prediction in PPMI showed significant but modest performance, with correlations of $r = 0.26$ for cross-sectional prediction and $r = 0.21$ for longitudinal prediction (both FDR-adjusted $P = 0.002$; Fig. 6C).

As a sensitivity analysis, we repeated external validation using direct external scoring without new-site recalibration. These analyses showed expected site-related shifts in HC centile distributions, but the major disease-related deviation patterns remained broadly consistent with the recalibrated analyses. The unrecalibrated external median centile profiles are provided in Fig. S3.

Added contribution of hippocampal subfield to the above clinical tasks against whole-brain structural metrics

To evaluate the added value of hippocampal subfield information in the above clinical tasks, we compared five feature sets under the corresponding downstream modeling frameworks. Whole-brain structural features were derived from the FreeSurfer-derived 82-ROI whole-brain model, comprising 68 Desikan–Killiany cortical ROIs and 14 subcortical ROIs and converted into centile scores using the same normative scoring framework. The five feature sets included bilateral whole-hippocampus centile scores only (M1), bilateral hippocampal subfield centile scores only (M2), combined bilateral whole-hippocampus and hippocampal subfield centile scores (M3; primary hippocampal model), whole-brain ROI centile scores including bilateral whole-hippocampal volumes (M4), and whole-brain ROI centile scores plus hippocampal subfield centile scores (M5). These comparisons were designed to assess the added value of subfields beyond whole-hippocampal measures alone (M3 vs. M1), the performance of hippocampus-focused features relative to whole-brain structural metrics (M3 vs. M4), and the incremental value of adding hippocampal subfields to whole-brain models (M5 vs. M4). Classification, cognitive prediction, and progression-related tasks were compared using the same task-specific modeling strategies and performance metrics as in the primary analyses. Detailed performance metrics are provided in Table S2.

For HC-versus-disease classification, M3 consistently outperformed M1 across all disease groups, with AUCs increasing from 0.55 to 0.89 for MCI, 0.78 to 0.86 for AD, 0.50 to 0.74 for PD, 0.53 to 0.80 for CSVD, and 0.61 to 0.91 for MS. Compared with M4, M3 achieved comparable or better performance in selected tasks, including MCI classification. Adding hippocampal subfield features to the whole-brain model further improved classification performance across all HC-versus-disease tasks, with M5 exceeding M4 by AUC increments of 0.01–0.06.

For cross-sectional cognitive prediction, M3 improved performance over M1 across multiple cognitive outcomes, indicating that hippocampal subfield centile scores provided additional information beyond bilateral whole-hippocampal measures alone. M3 also showed comparable or superior performance to M4 in several tasks, particularly MCI cognitive prediction and MS-related cognitive prediction. Adding hippocampal subfield features to whole-brain ROI features further improved prediction performance for several outcomes, including MCI-MMSE (M5 vs. M4, $\Delta r = 0.110$) and MS-SDMT ($\Delta r = 0.097$).

For longitudinal cognitive prediction and disease progression analyses, M3 continued to improve over M1 in several tasks. M3 was also comparable to or better than M4 for MCI-to-AD conversion prediction. Adding hippocampal subfield features to the whole-brain model produced the largest incremental gain for MCI-to-AD conversion prediction (M5 vs. M4, ΔC-index = 0.152), whereas gains were smaller or absent for some longitudinal cognitive prediction and MS progression outcomes.

Overall, M5 improved or maintained performance relative to M4 across all 26 benchmark comparisons, with improvements in 22 tasks and unchanged performance in 4 tasks. These findings indicate that hippocampal subfield centile scores provide clinically relevant information beyond bilateral whole-hippocampal measures alone and can add incremental value to whole-brain structural models across classification, cognitive prediction, and progression-related tasks.

Clinically applicable personalized hippocampal health report

To facilitate individualized interpretation of hippocampal deviations, we developed a prototype personalized hippocampal health report. The report includes visual maps of hippocampal subfield deviations for rapid inspection, summaries of raw volumes and centile scores with values below the 2.5th centile explicitly flagged, DPS estimates, and predicted cognitive scores derived from the corresponding classification and prediction models. DPS estimates and predicted cognitive scores were generated from previously established disease-classification and cognitive-prediction models based on hippocampal centile scores.

This report is not intended to serve as a standalone diagnostic tool, but rather to provide complementary neuroanatomical information when MRI is clinically indicated. As an illustrative example, we present the report of one patient with AD in Fig. S8, demonstrating how individualized hippocampal and subfield deviations can be visualized and summarized at the single-patient level.

Sensitivity analyses

To evaluate whether the normative trajectories and downstream findings were robust to key modeling choices and potential sources of variability, we performed four sets of sensitivity analyses: (1) ICV-adjusted normative modeling and downstream validation, (2) age-restricted normative modeling, (3) scanner- and site-related robustness analyses, and (4) sex-stratified normative modeling.

First, ICV-adjusted models showed high agreement with the primary non-ICV-adjusted models. Lifespan trajectories were nearly identical, with a median trajectory similarity of $r = 0.999$, and individualized centile scores also showed high concordance with the primary model (median concordance correlation coefficient = 0.942). Repeating disease classification and cognitive prediction analyses using ICV-adjusted centile scores yielded performance broadly comparable to the primary analyses, indicating that the diagnostic and cognitive predictive utility of hippocampal centile scores was not driven solely by inter-individual differences in head size (Fig. S6).

Second, the age-restricted model fitted in participants aged 18–70 years showed highly similar normative trajectories to the primary model, with a median trajectory similarity of $r = 0.990$ and a median concordance correlation coefficient of 0.999 for individualized centile scores. This finding indicated that the main normative patterns were not driven by the youngest or oldest portions of the reference sample.

Third, scanner- and site-related sensitivity analyses demonstrated that the normative centile estimates were robust to nonlinear harmonization and site-level scale adjustment. After ComBat-GAM harmonization, individualized centile estimates remained highly consistent with those from the primary models across all 24 hippocampal regions, with Spearman correlation coefficients ranging from 0.9584 to 0.9979 and mean absolute centile differences ranging from 1.27 to 5.95 percentile points. Similarly, empirical site-specific scale adjustment had minimal impact on individualized centile estimates, with a mean Spearman correlation of approximately 0.999 and a mean absolute centile difference of approximately 1 centile point relative to the primary models (Table S3).

Finally, sex-stratified analyses showed broadly consistent disease-related deviation patterns relative to sex-specific normative references, with selected subfield-level sex differences observed across neurological disorders (Fig. S5).

Discussion

In this study, we established lifespan normative references for whole-hippocampal and hippocampal subfield volumes using 34,789 HCs aged 4–85 years from multiple international cohorts. These references captured nonlinear and regionally heterogeneous lifespan trajectories and enabled individualized centile scoring of hippocampal deviations. Applied to MCI, AD, PD, CSVD, and MS, the centile scores revealed disease-related deviation patterns that were associated with cognitive performance and longitudinal clinical outcomes. Independent validation in ADNI and PPMI supported the out-of-sample generalizability of the framework, while benchmark comparisons against whole-hippocampus and whole-brain models demonstrated the complementary value of hippocampal subfield information. Together, these findings suggest that hippocampal subfield centile scores provide interpretable individualized measures for characterizing hippocampal involvement across neurological disorders.

This study provides a comprehensive normative framework for hippocampal and subfield volumes by leveraging a large lifespan cohort and GAMLSS modeling. Consistent with established models of brain structural maturation and senescence, the whole hippocampus and most subfields followed nonlinear inverted U-shaped trajectories, with peak ages ranging from 27.4 to 40.1 years, providing developmentally informed references to distinguish normal aging from disease-related alterations[30,33]. The earliest post-childhood peak was observed in the fimbria, whereas the latest peak was observed in CA3, suggesting regionally heterogeneous maturational and aging profiles across hippocampal subfields[34–36]. The parasubiculum showed a near-linear decline from early life, consistent with previous reports of early maturation in this region[37,38]. Its connectivity and neurochemical features may contribute to its sensitivity to aging and disease-related processes. We also observed hemispheric asymmetries, with left subfields peaking earlier than right counterparts, potentially reflecting lateralized developmental or functional lateralization[39]. The estimated rates of change further indicated a transition from maturational growth to accelerating decline, aligning with established lifespan models[31]. These normative references highlight divergent developmental timelines across hippocampal subfields and provide a benchmark for interpreting subfield-specific alterations in aging and neurological disease[38,40].

Compared with matched HCs, all disease cohorts showed reductions in whole-hippocampal centile scores, together with heterogeneous subfield-level alterations[41]. Although reduced centile scores were observed across multiple disorders, the most affected subfields differed by diagnosis. These regions may be vulnerable to shared pathological processes, including neuroinflammation, vascular insufficiency, and protein aggregation[42–45]. Despite these shared vulnerabilities, group-average subfield deviation patterns differed across disease groups. MCI patients exhibited relatively mild but detectable subfield reductions, with the largest effects in the right parasubiculum, left presubiculum, and left CA4, which may reflect early hippocampal involvement along the Alzheimer's disease continuum[46,47]. However, because MCI encompasses clinically heterogeneous phenotypes, including amnestic and nonamnestic subtypes[48], group-level estimates may dilute subtype-specific hippocampal involvement and partly explain why CMD was not significantly increased in MCI after FDR correction. AD patients showed widespread reductions across all subfields, with the strongest effects in the bilateral subiculum and right molecular layer, consistent with the known vulnerability of hippocampal output and integrative regions in Alzheimer's disease[49,50]. PD showed more focal alterations, with the largest effects in the presubiculum and HATA, potentially reflecting disruption of hippocampal circuits involved in cognitive and limbic functions[6,41,51,52]. CSVD preferentially affected the hippocampal tail and presubiculum, which may relate to vascular vulnerability and the contribution of posterior hippocampal structures to memory and spatial processing[53,54]. MS showed widespread subfield reductions, with prominent involvement of the hippocampal tail and HATA, consistent with the contribution of neuroinflammatory and demyelinating processes to hippocampal injury[55,56]. Together, these findings indicate that subfield-level centile scores can reveal disease-specific hippocampal involvement beyond global hippocampal volume. As a complementary multivariate summary, CMD captured the overall burden of hippocampal deviation[31], with significant increases in AD, CSVD, and MS and the largest effect observed in MS. The lack of significant CMD differences in MCI and PD suggests that CMD may be most informative when hippocampal involvement is broader or more severe, whereas region-wise centile profiles remain necessary for capturing subtle, focal, or heterogeneous subfield alterations.

The disease-specific classification models based on hippocampal subfield centile scores achieved performance comparable to previous findings[51,57], indicating that individualized hippocampal deviation patterns contain disease-relevant information[48]. Furthermore, the DPS summarized multivariate hippocampal deviation patterns into an The DPS further summarized multivariate hippocampal deviation patterns into an interpretable model-derived index of similarity to disease-specific patterns[58]. In multi-disease contexts, such pattern-level measures may complement region-wise centile profiles for differential characterization and individualized stratification. Nonetheless, DPS estimates should be interpreted as supportive model-derived measures rather than standalone diagnostic probabilities, and broader prospective validation remains necessary to establish clinical utility.

Hippocampal and subfield centile scores were associated with cognitive performance, with disease-specific heterogeneity in the most relevant subfields. In MCI and AD, the strongest associations involved the presubiculum and fimbria, respectively, consistent with the role of hippocampal input-output pathways in episodic memory and global cognition[14]. In PD, cognitive associations were most evident in the molecular layer and GC-DG, whereas CSVD showed prominent associations involving the subiculum and GC-DG, regions implicated in memory encoding, retrieval, and hippocampal-cortical communication[7,51]. In MS, the strongest associations were observed in the left fimbria for SDMT, right CA4 for BVMT-R, and left subiculum for PASAT, suggesting that hippocampal subfield deviations may be relevant to multiple cognitive domains, including processing speed, visuospatial memory, and working memory[8]. These region- and disease-specific associations suggest that centile scores capture clinically meaningful structure-cognition relationships. When interpreted alongside cognitive assessment, hippocampal centile scores may provide complementary neurobiological information for understanding cognitive heterogeneity across disorders. Individuals with similar cognitive scores may exhibit different degrees of hippocampal structural deviation, suggesting distinct underlying neurobiological substrates. Thus, centile scores may help contextualize cognitive performance in relation to individualized hippocampal integrity, rather than merely serving as predictors of cognitive scores.

Our results further suggest that hippocampal subfield centile scores have predictive utility for longitudinal cognitive and clinical outcomes. Beyond cross-sectional cognitive associations, centile scores predicted follow-up cognitive performance in MCI, AD, and CSVD and supported risk stratification for HC-to-MCI conversion, MCI-to-AD conversion, EDSS worsening, and SPMS conversion. These findings are consistent with prior hippocampal research showing that structural alterations can stratify individuals by risk[59]. Subfield-specific deviations may therefore reflect disease-related hippocampal involvement that is not fully captured by global hippocampal measures[60]. By benchmarking individual hippocampal anatomy against normative references, centile scores may provide a scalable framework for individualized risk stratification and longitudinal interpretation[61].

The external validation analyses further support the generalizability of the proposed normative framework. In two independent

public datasets not used for model construction, new-site recalibrated centile scores showed appropriate calibration in external HCs and broadly reproduced disease-related hippocampal deviation patterns across the Alzheimer's disease continuum and Parkinson's disease[62]. These external analyses further supported disease discrimination, cognitive prediction, and prognostic stratification. Direct external scoring without recalibration showed expected site-related shifts in HC centile distributions, highlighting the importance of site calibration when applying normative models to previously unseen scanners or cohorts[63]. Together, these findings support the robustness and transportability of the framework while underscoring the need for further prospective validation across additional diseases, populations, and acquisition settings.

Benchmark analyses clarified the contribution of hippocampal subfield information relative to bilateral whole-hippocampal and whole-brain ROI measures[64]. The full hippocampal model consistently outperformed the whole-hippocampus-only model, supporting the added value of subfield-level centile scores beyond global hippocampal measures. Despite its lower dimensionality, the combined hippocampal model achieved performance comparable to or exceeding the whole-brain ROI model in several tasks, including MCI classification, MCI cognitive prediction, MS-related cognitive prediction, and MCI-to-AD conversion prediction. Moreover, adding hippocampal subfield features to the whole-brain ROI model improved or maintained performance across all benchmark comparisons, suggesting that hippocampal subfield centile scores provide compact, anatomically interpretable, and complementary information to whole-brain structural measures.

Finally, we developed a personalized hippocampal health report that integrates subfield volumes, centile scores, flagged infra-normal deviations, DPS estimates, and predicted cognitive scores. By translating multidimensional hippocampal deviation profiles into an interpretable single-patient summary, the report illustrates one potential route toward individualized interpretation of normative modeling results. This report may provide complementary neuroanatomical information to support individualized interpretation when MRI is clinically indicated.

This study has several limitations. First, several constraints should be considered in the construction and interpretation of the normative models. The models were built retrospectively using data from multiple sites, which may introduce heterogeneity related to participant selection, scanner vendor, field strength, and acquisition protocols. Although site effects were modeled and additional ComBat-GAM harmonization and site-scale sensitivity analyses supported the robustness of the centile estimates, residual scanner- or protocol-related variability may persist[65]. Moreover, the normative dataset did not include children younger than 4 years and had relatively sparse sampling in certain age ranges[66,67]. In addition, age-related change rates were inferred from cross-sectional rather than longitudinal trajectories. These factors may affect both the generalizability and the precision of the lifespan reference models. Second, hippocampal subfield estimates remain sensitive to segmentation variability. The small size and less distinct anatomical boundaries of several subfields make them prone to measurement error, and different segmentation tools, such as FreeSurfer, ASHS, and HippUnfold, can yield variable estimates[68]. Such methodological variability may limit the comparability of findings across studies and introduce additional uncertainty into deviation estimates. Third, several aspects of the patient data may limit clinical generalizability. Although the overall patient sample was substantial, some disease groups, clinical subgroups, and longitudinal subsets were relatively modest, and follow-up durations were limited, constraining the strength of prognostic inferences. Cognitive data were incomplete and largely restricted to brief global screening measures such as MMSE and MoCA or selected domain-specific tests, which provide relatively coarse estimates of cognitive function and limit comprehensive evaluation of structure-function relationships. In addition, disease subtypes were not explicitly modeled. This limitation is particularly relevant for MCI, which encompasses amnestic and nonamnestic as well as single-domain and multidomain subtypes that may have distinct hippocampal involvement[48]. Because detailed MCI subtype labels were not consistently available in the primary cohort, we could not directly evaluate subtype-specific hippocampal deviation profiles. Aggregating heterogeneous subgroups may therefore have produced group-averaged deviation patterns and influenced the performance of some disease-specific models.

Future research could extend hippocampal subfield normative models by integrating them with whole-brain normative frameworks, large-scale brain networks, and complementary imaging modalities. Such integrated approaches may capture both focal hippocampal abnormalities and distributed brain alterations, thereby improving disease classification, risk stratification, and individualized clinical interpretation. In addition, large-scale, multi-ethnic, prospective longitudinal cohorts with standardized imaging protocols and comprehensive cognitive assessments will be crucial for validating these models, refining their generalizability, and advancing their clinical translation.

In conclusion, we established lifespan normative references for whole-hippocampal and hippocampal subfield volumes using 3D T1w-MRI from 34,789 healthy controls aged 4–85 years. From these references, we derived personalized centile scores that provide a standardized framework for assessing hippocampal deviations across neurological disorders, capturing disease-specific patterns, providing complementary neurobiological information for interpreting cognitive heterogeneity, and supporting prediction of longitudinal clinical outcomes. Furthermore, prototype personalized hippocampal health reports may offer an interpretable format for summarizing individual hippocampal deviation profiles and supporting future clinical translation.

Methods

Study design

This was a large-scale, multicenter neuroimaging study combining cross-sectional and longitudinal MRI and clinical data. The primary aim was to construct lifespan normative references for whole-hippocampal and hippocampal subfield volumes using 3D T1w-MRI from 34,789 healthy controls aged 4–85 years from multiple international cohorts. Individualized hippocampal deviation scores were then derived as centile scores relative to the normative references.

These centile scores were applied to patients with neurological disorders, including MCI, AD, PD, CSVD, and MS, to characterize disease-related hippocampal deviation patterns and evaluate their associations with cognitive performance, disease classification, cognitive prediction, and longitudinal outcomes. To assess out-of-sample generalizability, we performed independent external validation using ADNI for the Alzheimer's disease continuum and PPMI for Parkinson's disease. To contextualize the utility of hippocampal subfield measures, we systematically compared hippocampal models with whole-hippocampus and whole-brain benchmark models under the same downstream modeling framework. Robustness was evaluated through sensitivity analyses described below. Finally, we developed an illustrative individualized visualization report to summarize hippocampal and subfield deviations at the single-participant level.

This study included both locally collected cohorts and publicly available and third-party research cohorts. The locally collected cohorts were approved by the Ethics Committees of Beijing Tiantan Hospital, Capital Medical University (Nos. KY-2019-050-02, KY-2019-140-02, KY-2024-221-02, and KY2025-203-03). Written informed consent was obtained from participants or their legal guardians where required, while the requirement for informed consent was waived by the relevant ethics committees for certain retrospective cohorts.

Publicly available and third-party cohorts were collected under the ethics approvals and informed-consent procedures of their respective original studies. No additional participant compensation was provided for the present secondary analyses. The study adhered to the STROBE guidelines.

Participants

Participants were assembled from multiple international cohorts. The normative reference sample included 34,789 healthy controls aged 4–85 years from the Chinese Brain Health Imaging Consortium (CHART), Information eXtraction from Images (IXI), Cambridge Centre for Ageing and Neuroscience (CamCAN), Southwest University Adult Lifespan Dataset (SALD), Strategic Research Program for Brain Sciences (SRPBS), Autism Brain Imaging Data Exchange (ABIDE), ADHD-200, Nathan Kline Institute–Rockland Sample (NKI-RS), Open Access Series of Imaging Studies (OASIS), and curated OpenNeuro cohorts. Cohort characteristics are provided in Table S1, and the normative reference age distribution is shown in Fig. S1. Patient analyses included 3,521 individuals with neurological disorders, including 308 patients with MCI, 604 with AD, 1226 with PD, 864 with CSVD, and 519 with MS. The inclusion criteria were as follows: (1) the availability of high-quality 3D T1w-MRI data; and (2) participant diagnoses based on internationally recognized clinical standards. HCs were defined as individuals between 4 and 85 years of age with no history of clinically diagnosed central nervous system (CNS) disorders. All patient diagnoses were made in accordance with internationally recognized clinical standards: AD was diagnosed according to the National Institute on Aging-Alzheimer's Association (NIA-AA) criteria[69]; idiopathic PD was defined based on the International Parkinson and Movement Disorder Society clinical diagnostic criteria[70]; CSVD was identified according to the STandards for ReportIng Vascular changes on nEuroimaging (STRIVE) guidelines[71]; and MS was classified as relapsing-remitting MS under the 2017 McDonald criteria, with confirmed negativity for both aquaporin-4 and myelin oligodendrocyte glycoprotein antibodies[72]. The exclusion criteria were as follows: (1) a history of other neurological or neuropsychiatric conditions (e.g., stroke or tumor); (2) incomplete MRI acquisition, severe motion artifacts, or poor image quality as assessed during visual inspection; and (3) failed segmentation or Euler index values more than 2 SD below the population median in FreeSurfer outputs[73].

Demographic information, including age, sex, and education, was obtained from the contributing cohorts. Demographic characteristics of the primary study groups, including age and sex distributions, are summarized in Table 1. Clinical evaluations were performed by qualified neurologists. Global cognitive function was assessed using the Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA) in HCs and in patients with MCI, AD, PD, and CSVD. In contrast, cognitive assessment in MS focused on information processing speed and memory using the Symbol Digit Modalities Test (SDMT), Brief Visuospatial Memory Test-Revised (BVMT-R), and Paced Auditory Serial Addition Test (PASAT).

Longitudinal MRI and cognitive assessments were available in subsets of HCs, MCI, AD, and CSVD participants, with follow-up durations ranging from 0.5 to 3 years and up to four time points. During follow-up, a subset of HCs converted to MCI and a subset of MCI participants progressed to AD. For MS, extended follow-up assessments included Expanded Disability Status Scale (EDSS) progression and conversion from relapsing-remitting MS to secondary progressive MS (SPMS)[74].

MRI acquisition and image processing

3D T1w-MRI scans were included from large-scale international cohorts and independent external validation datasets. Cohort-specific imaging sources and acquisition parameters for the healthy control cohorts incorporated into the expanded normative reference are provided in Table S1. Quality control followed a two-step procedure: (1) independent visual inspection by trained raters, with discrepancies resolved by a senior neuroradiologist; and (2) automated assessment using the Euler number, with outliers defined as values at least two standard deviations below the population median.

Image preprocessing, including intensity normalization, skull stripping, and Talairach registration, was performed using FreeSurfer version 7.3.2[75]. The segmentation workflow involved the following two steps. First, the bilateral hippocampi were delineated to obtain total hippocampal volumes. Second, each hippocampus was further segmented into 19 anatomical subfields: the parasubiculum, presubiculum-head, presubiculum-body, subiculum-head, subiculum-body, CA1-head, CA1-body, CA3-head, CA3-body, CA4-head, CA4-body, GC-DG-head, GC-DG-body, molecular layer–HP-head, molecular layer-HP-body, HATA, fimbria, tail, and fissure. For subsequent analyses, these 19 subfields were consolidated following the FS60 scheme, yielding 11 volumetric features: CA1, CA3, CA4, GC-DG, ML, presubiculum, subiculum, parasubiculum, HATA, fimbria, and tail. The hippocampal fissure was excluded from analysis. Raw volumes (mm^3) for the left and right hippocampus were extracted, visually inspected, and manually edited when necessary.

Normative modeling

Normative lifespan models for hippocampal and hippocampal subfield volumes were constructed using GAMLSS in R and trained exclusively on the HC dataset ($n = 34{,}789$). Age and sex were modeled as fixed effects, and site was included as a random effect. A generalized gamma distribution was used to accommodate skewed volumetric distributions; age effects were modeled using fractional polynomials, and model complexity was selected according to the Bayesian information criterion (BIC). Tenfold cross-validation was used to derive out-of-fold centile scores for HCs, ensuring that each HC was scored using a model that did not include that individual. For patients, centile scores were computed by applying the final HC-trained normative models. Personalized centile scores were derived from the fitted age-, sex-, and site-adjusted normative distributions using the cumulative distribution function of the model. All patient scans originated from sites represented in the normative dataset; scoring procedures for previously unseen external sites are described in the external validation analysis. To quantify uncertainty in centile curves and peak ages, we performed subject-level bootstrap resampling with replacement over 1000 iterations and refitted the model in each resample. Model-derived absolute and relative rates of change were computed from the first derivative of the fitted median trajectory for each structure. Absolute rates were expressed in mm^3/year, whereas relative rates were normalized by the model-predicted median volume and expressed as %/year.

Independent external validation

To evaluate out-of-sample generalizability, we performed external validation using two independent public datasets that were not included in normative model construction: ADNI for the Alzheimer's disease continuum and PPMI for Parkinson's disease. The external cohorts included HCs, patients with MCI, and patients with AD from ADNI, as well as HCs and patients with PD from PPMI. All MRI scans were processed using the same FreeSurfer pipeline as the primary cohort.

Participants were eligible if they had baseline T1w MPRAGE/MP-RAGE MRI scans, available demographic information, and successful image quality control. To reduce diagnostic imbalance and ensure comparable group sizes, we used a prespecified balanced sampling strategy after quality filtering. Specifically, 200 participants were randomly selected from each ADNI diagnostic group (HC, MCI, and AD), and 200 HCs and 200 patients with PD were randomly selected from PPMI. Participants were selected before normative scoring and

downstream model evaluation, and no participant was selected based on centile scores, disease-related deviation patterns, or classification performance. The normative models were not refitted in the external datasets. External participants were scored using the normative models derived from the primary healthy reference cohort. Because ADNI and PPMI originated from previously unseen sites and scanners, we performed new-site recalibration using external HCs to reduce site-related offsets while preserving the original normative age- and sex-related trajectories. Disease-related analyses were then conducted using the recalibrated centile scores. External validation assessed centile calibration in HCs, disease-related deviation patterns, and downstream clinical utility. Specifically, we evaluated whether external HCs showed approximately centered centile distributions, whether disease-related hippocampal subfield deviation patterns were directionally consistent with the primary cohort, and whether MCI and AD showed graded hippocampal deviations. In ADNI, we additionally compared EMCI and LMCI to assess stage-related heterogeneity of hippocampal centile deviations within the MCI spectrum. We further evaluated external classification and cognitive prediction using the same downstream analytic frameworks as the primary analyses. In ADNI, we also assessed prognostic utility for clinical progression. As a sensitivity analysis, external validation was repeated without new-site recalibration to evaluate robustness under direct out-of-sample application.

Benchmark comparison of demographic, hippocampal, and whole-brain feature sets across downstream clinical tasks

To evaluate the added value of imaging-derived hippocampal subfield information in the above clinical tasks, we constructed six feature sets for downstream analyses. M0 included age and sex only and served as a demographic baseline model where applicable. M1 included bilateral whole-hippocampal centile scores only. M2 included bilateral hippocampal subfield centile scores only. M3 included both bilateral whole-hippocampal and hippocampal subfield centile scores and was defined as the primary hippocampal model. M4 included whole-brain ROI centile scores derived from the FreeSurfer-derived 82-ROI whole-brain model, comprising 68 Desikan–Killiany cortical ROIs and 14 subcortical ROIs, including bilateral whole-hippocampal volumes. M5 combined whole-brain ROI centile scores with bilateral hippocampal subfield centile scores. These feature sets were designed to assess the contribution of demographic variables alone (M0), the added value of subfields beyond whole-hippocampal measures alone (M3 vs. M1), the performance of hippocampus-focused features relative to whole-brain structural metrics (M3 vs. M4), and the incremental value of adding hippocampal subfields to whole-brain models (M5 vs. M4). All feature sets were evaluated using the same task-specific downstream modeling frameworks and performance metrics as in the primary analyses.

Statistical analysis

All statistical analyses were performed in R (version 4.4.1). Detailed software packages and version information are provided in Supplementary Methods S6. Categorical variables were summarized as counts and percentages, and continuous variables as mean ± standard deviation (SD) or median [interquartile range (IQR)], as appropriate. Descriptive group comparisons used Pearson's chi-square tests, one-way analysis of variance, or Kruskal–Wallis tests, as appropriate. All downstream analyses used precomputed normative centile scores as fixed input features. Group differences in hippocampal and hippocampal subfield centile scores were assessed using 1:1 matched Monte Carlo permutation tests based on Welch's t statistic. For each disease-versus-HC comparison, cases were matched without replacement to HCs from the same site/scanner and sex using nearest-age matching, and diagnostic labels were permuted within matched pairs. Two-sided permutation P values were estimated using 5000 permutations with finite-sample correction. Effect sizes were reported as Cohen's d with 95% confidence intervals, and regional P values were adjusted across the 24 hippocampal ROIs within each disease group using the Benjamini–Hochberg false discovery rate (FDR) procedure. CMD was calculated from the complete 24-dimensional hippocampal centile profile as detailed in Supplementary Method S5 and analyzed using the same matched permutation framework, with FDR correction across disease groups. Associations of hippocampal centile scores and CMD with cognitive measures were assessed using partial Spearman correlations controlling for age and sex. Regional association P values were FDR-adjusted across the 24 hippocampal ROIs separately within each analysis-group-by-cognitive-outcome combination, whereas CMD associations were corrected separately across cognitive outcomes within each analysis group. For disease classification, the eligible dataset for each HC-versus-disease task was divided into training and independent held-out test sets using a stratified 7:3 split. Within each partition, cases were matched 1:1 without replacement to same-sex HCs using nearest-age matching. Regularized logistic regression and single-hidden-layer neural network classifiers were tuned within the training set using 5-fold cross-validation, and the model with the highest cross-validated ROC AUC was evaluated on the held-out test set. AUCs were calculated from held-out predicted probabilities, with 95% confidence intervals estimated using DeLong's method. Disease propensity scores were defined as the held-out predicted probabilities for the target disease. Cognitive prediction used radial-basis-function support vector regression with outer 10-fold cross-validation and 5-fold internal tuning. Age, sex, and the 24 hippocampal centile features were included as predictors, with preprocessing and model fitting performed using training data only. Longitudinal observations from the same participant were assigned to the same outer fold. Prediction performance was quantified using Pearson's correlation between observed and predicted scores, with Benjamini–Hochberg correction applied separately across all cross-sectional tasks and all longitudinal tasks. Prognostic analyses used outer 10-fold cross-validation with 5-fold internal tuning of LASSO Cox models. Out-of-fold risk scores were used to calculate Harrell's concordance index and to define high- and low-risk groups for Kaplan–Meier and two-sided log-rank analyses. Missing data were handled using analysis-specific complete-case analyses without imputation. All statistical tests were two-sided, with statistical significance set at $P < 0.05$ unless otherwise specified.

Sensitivity analyses

To evaluate the robustness of the normative models and downstream findings, we conducted several sensitivity analyses. First, we constructed ICV-adjusted normative models by including intracranial volume as an additional fixed-effect covariate. Agreement between the ICV-adjusted and primary non-ICV-adjusted models was assessed by comparing lifespan trajectories and individualized centile scores. We further repeated disease classification and cognitive prediction analyses using ICV-adjusted centile scores and compared performance with the primary analyses using the same training-test partitions or cross-validation folds, as appropriate, together with the same performance metrics. Second, we examined the influence of age range by refitting the normative models after restricting the reference sample to individuals aged 18–70 years and comparing the resulting trajectories and individualized centile scores with those from the primary models. Third, we assessed scanner- and site-related robustness using nonlinear ComBat-GAM harmonization and an empirical site-specific scale sensitivity analysis; detailed procedures are provided in Method S2. Finally, we assessed sex-related effects by fitting sex-stratified normative models and comparing disease-related deviations relative to sex-specific references.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Individual-level MRI and clinical data are not publicly available and cannot be redistributed by the authors because access to and redistribution of these data are governed by the original institutional approvals, participant consent, and data-use agreements. Data from third-party cohorts remain available through their respective data-access procedures. The normative reference resources and example data used in this study are publicly available at GitHub: https://github.com/xiaozhoudu-lab/normative-hippocampal-subfields. Source data are provided with this paper.

Code availability

All code used for normative modeling, centile scoring, statistical analyses, and figures is publicly available at GitHub: https://github.com/xiaozhoudu-lab/normative-hippocampal-subfields

References

1. Liu, N. et al. Cross-ancestry genome-wide association meta-analyses of hippocampal and subfield volumes. *Nat. Genet.* **55**, 1126–1137 https://doi.org/10.1038/s41588-023-01425-8 (2023).
2. Touron, E. et al. Depressive symptoms in older adults are associated with changes in stress-related markers, functional connectivity and brain volume. *Alzheimer's Res. Ther.* **17**, 9 https://doi.org/10.1186/s13195-024-01643-0 (2025).
3. Ballesteros, C. et al. Linking birth experience and perinatal depression symptoms to neuroanatomical changes in hippocampus and amygdala. *Sci. Adv.* **11**, eadt5619 https://doi.org/10.1126/sciadv.adt5619 (2025).
4. Zilioli, A. et al. Unveiling the hippocampal subfield changes across the Alzheimer's disease continuum: a systematic review of neuroimaging studies. *Brain Imaging Behav.* **19**, 253–267 https://doi.org/10.1007/s11682-024-00952-0 (2025).
5. Grabrucker, S. et al. Microbiota from Alzheimer's patients induce deficits in cognition and hippocampal neurogenesis. *Brain* **146**, 4916–4934 https://doi.org/10.1093/brain/awad303 (2023).
6. Belloso-Iguerategui, A. et al. Hippocampal synaptic failure is an early event in experimental parkinsonism with subtle cognitive deficit. *Brain* **146**, 4949–4963 https://doi.org/10.1093/brain/awad227 (2023).
7. Perosa, V. et al. Association between hippocampal volumes and cognition in cerebral amyloid angiopathy. *Neurology* **102**, e207854 https://doi.org/10.1212/wnl.0000000000207854 (2024).
8. Rocca, M. A. et al. Regional hippocampal involvement and cognitive impairment in pediatric multiple sclerosis. *Mult. Scler.* **22**, 628–640 https://doi.org/10.1177/1352458515598569 (2016).
9. Morrison, C., Dadar, M., Shafiee, N. & Collins, D. L. The use of hippocampal grading as a biomarker for preclinical and prodromal Alzheimer's disease. *Hum. Brain Mapp.* **44**, 3147–3157 https://doi.org/10.1002/hbm.26269 (2023).
10. Dimsdale-Zucker, H. R., Ritchey, M., Ekstrom, A. D., Yonelinas, A. P. & Ranganath, C. CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. *Nat. Commun.* **9**, 294 https://doi.org/10.1038/s41467-017-02752-1 (2018).
11. Punzi, M. et al. Atrophy of hippocampal subfields and amygdala nuclei in subjects with mild cognitive impairment progressing to Alzheimer's disease. *Heliyon* **10**, e27429 https://doi.org/10.1016/j.heliyon.2024.e27429 (2024).
12. Constantinides, V. C. et al. Hippocampal subfield volumetry in corticobasal syndrome of diverse underlying pathologies. *J. Neurol.* **270**, 2059–2068 https://doi.org/10.1007/s00415-022-11538-5 (2023).
13. Iglesias, J. E. et al. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. *Neuroimage* **115**, 117–137 https://doi.org/10.1016/j.neuroimage.2015.04.042 (2015).
14. Zhang, J. et al. Hippocampal subfield volumes in mild cognitive impairment and alzheimer's disease: a systematic review and meta-analysis. *Brain Imaging Behav.* **17**, 778–793 https://doi.org/10.1007/s11682-023-00804-3 (2023).
15. Becker, S. et al. Association of hippocampal subfields, CSF biomarkers, and cognition in patients with Parkinson disease without dementia. *Neurology* **96**, e904–e915 https://doi.org/10.1212/wnl.0000000000011224 (2021).
16. Cortese, R. et al. Regional hippocampal atrophy reflects memory impairment in patients with early relapsing remitting multiple sclerosis. *J. Neurol.* https://doi.org/10.1007/s00415-024-12290-8 (2024).
17. Cacciaguerra, L. et al. Volume of hippocampal subfields and cognitive deficits in neuromyelitis optica spectrum disorders. *Eur. J. Neurol.* **28**, 4167–4177 https://doi.org/10.1111/ene.15073 (2021).
18. Lotan, I. et al. Volumetric brain changes in MOGAD: a cross-sectional and longitudinal comparative analysis. *Mult. Scler. Relat. Disord.* **69**, 104436 https://doi.org/10.1016/j.msard.2022.104436 (2023).
19. Risacher, S. L. & Saykin, A. J. Neuroimaging in aging and neurologic diseases. *Handb. Clin. Neurol.* **167**, 191–227 https://doi.org/10.1016/b978-0-12-804766-8.00012-1 (2019).
20. Lewis, A. S. & Nobis, W. P. Narrowing the translational research gap by aligning replication concepts in basic and clinical neuroscience. *J. Neurosci. Methods* **392**, 109878 https://doi.org/10.1016/j.jneumeth.2023.109878 (2023).
21. Botvinik-Nezer, R. & Wager, T. D. Reproducibility in neuroimaging analysis: challenges and solutions. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **8**, 780–788 https://doi.org/10.1016/j.bpsc.2022.12.006 (2023).
22. Young, P. N. E. et al. Imaging biomarkers in neurodegeneration: current and future practices. *Alzheimer's Res. Ther.* **12**, 49 https://doi.org/10.1186/s13195-020-00612-7 (2020).
23. Pini, L. et al. Brain atrophy in Alzheimer's disease and aging. *Ageing Res. Rev.* **30**, 25–48 https://doi.org/10.1016/j.arr.2016.01.002 (2016).
24. Fletcher, E. et al. Brain volume change and cognitive trajectories in aging. *Neuropsychology* **32**, 436–449 https://doi.org/10.1037/neu0000447 (2018).
25. MacDonald, M. E. & Pike, G. B. MRI of healthy brain aging: a review. *NMR Biomed.* **34**, e4564 https://doi.org/10.1002/nbm.4564 (2021).
26. Rutherford, S. et al. Evidence for embracing normative modeling. *Elife* **12**, https://doi.org/10.7554/eLife.85082 (2023).
27. Qian, X. & Zhou, J. H. Imaging lifespan brain structural growth: from region, to connectome, to gradient. *PLoS Biol.* **22**, e3002669 https://doi.org/10.1371/journal.pbio.3002669 (2024).
28. Worker, A. et al. Extreme deviations from the normative model reveal cortical heterogeneity and associations with negative symptom severity in first-episode psychosis from the OPTiMiSE and GAP studies. *Transl. Psychiatry* **13**, 373 https://doi.org/10.1038/s41398-023-02661-6 (2023).
29. Verdi, S. et al. Personalizing progressive changes to brain structure in Alzheimer's disease using normative modeling. *Alzheimer's Dement.* **20**, 6998–7012 https://doi.org/10.1002/alz.14174 (2024).
30. Yu, J. Age-related decline in thickness and surface area in the cortical surface and hippocampus: lifespan trajectories and decade-by-decade analyses. *Geroscience* **46**, 6213–6227 https://doi.org/10.1007/s11357-024-01220-1 (2024).
31. Bethlehem, R. A. I. et al. Brain charts for the human lifespan. *Nature* **604**, 525–533 https://doi.org/10.1038/s41586-022-04554-y (2022).
32. Zhuo, Z. et al. Charting brain morphology in international healthy and neurological populations. *Nat. Neurosci.* **29**, 420–434 https://doi.org/10.1038/s41593-025-02144-5 (2026).
33. Fujita, S. et al. Characterization of brain volume changes in aging individuals with normal cognition using serial magnetic resonance

imaging. *JAMA Netw. Open* **6**, e2318153 https://doi.org/10.1001/jamanetworkopen.2023.18153 (2023).
34. Jessberger, S. & Gage, F. H. Stem-cell-associated structural and functional plasticity in the aging hippocampus. *Psychol. Aging* **23**, 684–691 https://doi.org/10.1037/a0014188 (2008).
35. Homayouni, R. et al. Age-related differences in hippocampal subfield volumes across the human lifespan: a meta-analysis. *Hippocampus* **33**, 1292–1315 https://doi.org/10.1002/hipo.23582 (2023).
36. Dohm-Hansen, S. et al. The 'middle-aging' brain. *Trends Neurosci.* **47**, 259–272 https://doi.org/10.1016/j.tins.2024.02.001 (2024).
37. Lavenex, P. & Banta Lavenex, P. Building hippocampal circuits to learn and remember: insights into the development of human memory. *Behav. Brain Res.* **254**, 8–21 https://doi.org/10.1016/j.bbr.2013.02.007 (2013).
38. Tamnes, C. K., Bos, M. G. N., van de Kamp, F. C., Peters, S. & Crone, E. A. Longitudinal development of hippocampal subregions from childhood to adulthood. *Dev. Cogn. Neurosci.* **30**, 212–222 https://doi.org/10.1016/j.dcn.2018.03.009 (2018).
39. Aumont, E. et al. Hippocampal subfield associations with memory depend on stimulus modality and retrieval mode. *Brain Commun.* **5**, fcad309 https://doi.org/10.1093/braincomms/fcad309 (2023).
40. Zheng, F. et al. The volume of hippocampal subfields in relation to decline of memory recall across the adult lifespan. *Front. Aging Neurosci.* **10**, 320 https://doi.org/10.3389/fnagi.2018.00320 (2018).
41. Bouwman, M. M. A. et al. Hippocampal subfields: volume, neuropathological vulnerability and cognitive decline in Alzheimer's and Parkinson's disease. *Alzheimer's Res. Ther.* **17**, 121 https://doi.org/10.1186/s13195-025-01768-w (2025).
42. Bussy, A. et al. Hippocampal subfield volumes across the healthy lifespan and the effects of MR sequence on estimates. *Neuroimage* **233**, 117931 https://doi.org/10.1016/j.neuroimage.2021.117931 (2021).
43. Muzio, L., Viotti, A. & Martino, G. Microglia in neuroinflammation and neurodegeneration: from understanding to therapy. *Front. Neurosci.* **15**, 742065 https://doi.org/10.3389/fnins.2021.742065 (2021).
44. Adamu, A., Li, S., Gao, F. & Xue, G. The role of neuroinflammation in neurodegenerative diseases: current understanding and future therapeutic targets. *Front. Aging Neurosci.* **16**, 1347987 https://doi.org/10.3389/fnagi.2024.1347987 (2024).
45. McCrea, M. et al. Mesoscale connectivity of the human hippocampus and fimbria revealed by ex vivo diffusion MRI. *Neuroimage* **310**, 121125 https://doi.org/10.1016/j.neuroimage.2025.121125 (2025).
46. Qu, H., Ge, H., Wang, L., Wang, W. & Hu, C. Volume changes of hippocampal and amygdala subfields in patients with mild cognitive impairment and Alzheimer's disease. *Acta Neurol. Belg.* **123**, 1381–1393 https://doi.org/10.1007/s13760-023-02235-9 (2023).
47. Xu, J., Tan, S., Wen, J., Zhang, M. & Xu, X. Progression of hippocampal subfield atrophy and asymmetry in Alzheimer's disease. *Eur. J. Neurosci.* **60**, 6091–6106 https://doi.org/10.1111/ejn.16543 (2024).
48. Emmert, N. A. et al. Hippocampal volumes in amnestic and non-amnestic mild cognitive impairment types using two common methods of MCI classification. *J. Int. Neuropsychol. Soc.* **28**, 391–400 https://doi.org/10.1017/s1355617721000564 (2022).
49. Ferreira, D. et al. The hippocampal sparing subtype of Alzheimer's disease assessed in neuropathology and in vivo tau positron emission tomography: a systematic review. *Acta Neuropathol. Commun.* **10**, 166 https://doi.org/10.1186/s40478-022-01471-z (2022).
50. Chen, Y. et al. Abnormal white matter changes in Alzheimer's disease based on diffusion tensor imaging: a systematic review. *Ageing Res. Rev.* **87**, 101911 https://doi.org/10.1016/j.arr.2023.101911 (2023).
51. Rau, A. et al. Hippocampus and basal forebrain degeneration differentially impact cognition in Lewy body spectrum disorders. *Brain* https://doi.org/10.1093/brain/awaf070 (2025).
52. Berlot, R. et al. Cholinergic basal forebrain and hippocampal structure influence visuospatial memory in Parkinson's disease. *Brain Imaging Behav.* **16**, 118–129 https://doi.org/10.1007/s11682-021-00481-0 (2022).
53. Haast, R. A. M. et al. Insights into hippocampal perfusion using high-resolution, multi-modal 7T MRI. *Proc. Natl. Acad. Sci. USA* **121**, e2310044121 https://doi.org/10.1073/pnas.2310044121 (2024).
54. Perosa, V. et al. Hippocampal vascular reserve associated with cognitive performance and hippocampal volume. *Brain* **143**, 622–634 https://doi.org/10.1093/brain/awz383 (2020).
55. Vlahovic, L. et al. Prevalence, demographic, and clinical factors associated with cognitive dysfunction in patients with neuromyelitis optica spectrum disorder. *Neurology* **102**, e207965 https://doi.org/10.1212/wnl.0000000000207965 (2024).
56. De Meo, E. et al. Effect of BDNF Val66Met polymorphism on hippocampal subfields in multiple sclerosis patients. *Mol. Psychiatry* **27**, 1010–1019 https://doi.org/10.1038/s41380-021-01345-1 (2022).
57. Ye, R. et al. Differential vulnerability of hippocampal subfields to amyloid and tau deposition in the Lewy body diseases. *Neurology* **102**, e209460 https://doi.org/10.1212/wnl.0000000000209460 (2024).
58. Ding, X. et al. Diabetes accelerates Alzheimer's disease progression in the first year post mild cognitive impairment diagnosis. *Alzheimer's Dement.* **20**, 4583–4593 https://doi.org/10.1002/alz.13882 (2024).
59. Broadhouse, K. M. et al. Hippocampal plasticity underpins long-term cognitive gains from resistance exercise in MCI. *Neuroimage Clin.* **25**, 102182 https://doi.org/10.1016/j.nicl.2020.102182 (2020).
60. Chen, Y. et al. Integrated cerebellar radiomic-network model for predicting mild cognitive impairment in Alzheimer's disease. *Alzheimer's Dement.* **21**, e14361 https://doi.org/10.1002/alz.14361 (2025).
61. Doran, S. et al. Relationship between hippocampal subfield volumes and cognitive decline in healthy subjects. *Front. Aging Neurosci.* **15**, 1284619 https://doi.org/10.3389/fnagi.2023.1284619 (2023).
62. Arani, A. et al. Design and validation of the ADNI MR protocol. *Alzheimer's Dement.* **20**, 6615–6621 https://doi.org/10.1002/alz.14162 (2024).
63. Fortin, J. P. et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage* **167**, 104–120 https://doi.org/10.1016/j.neuroimage.2017.11.024 (2018).
64. Parker, T. D. et al. Generalizable MRI normative modelling to detect age-inappropriate neurodegeneration. *Alzheimer's Res. Ther.* **17**, 244 https://doi.org/10.1186/s13195-025-01872-x (2025).
65. Reynolds, M., Chaudhary, T., Eshaghzadeh Torbati, M., Tudorascu, D. L. & Batmanghelich, K. ComBat harmonization: empirical Bayes versus fully Bayes approaches. *Neuroimage Clin.* **39**, 103472 https://doi.org/10.1016/j.nicl.2023.103472 (2023).
66. Sorrells, S. F. et al. Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. *Nature* **555**, 377–381 https://doi.org/10.1038/nature25975 (2018).
67. Kozareva, D. A., Cryan, J. F. & Nolan, Y. M. Born this way: hippocampal neurogenesis across the lifespan. *Aging Cell* **18**, e13007 https://doi.org/10.1111/acel.13007 (2019).
68. Mazloum-Farzaghi, N., Barense, M. D., Ryan, J. D., Stark, C. E. L. & Olsen, R. K. The effect of segmentation method on medial temporal lobe subregion volumes in aging. *Hum. Brain Mapp.* **45**, e70054 https://doi.org/10.1002/hbm.70054 (2024).
69. McKhann, G. M. et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* **7**, 263–269 https://doi.org/10.1016/j.jalz.2011.03.005 (2011).
70. Postuma, R. B. et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord.* **30**, 1591–1601 https://doi.org/10.1002/mds.26424 (2015).
71. Wardlaw, J. M. et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* **12**, 822–838 https://doi.org/10.1016/s1474-4422(13)70124-8 (2013).

72. Thompson, A. J. et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173 https://doi.org/10.1016/s1474-4422(17)30470-2 (2018).
73. Rosen, A. F. G. et al. Quantitative assessment of structural image quality. *Neuroimage* **169**, 407–418 https://doi.org/10.1016/j.neuroimage.2017.12.059 (2018).
74. Zhuo, Z. et al. Subtyping relapsing-remitting multiple sclerosis using structural MRI. *J. Neurol.* **268**, 1808–1817 https://doi.org/10.1007/s00415-020-10376-7 (2021).
75. Sämann, P. G. et al. FreeSurfer-based segmentation of hippocampal subfields: a review of methods and applications, with a novel quality control procedure for ENIGMA studies and other collaborative efforts. *Hum. Brain Mapp.* **43**, 207–233 https://doi.org/10.1002/hbm.25326 (2022).

Acknowledgements

We sincerely thank all participants, investigators, clinicians, research staff, imaging technicians, and site contributors involved in the CHART consortium and all public datasets used in this study. Data used in the preparation of this article were obtained from multiple public neuroimaging resources, including the Information eXtraction from Images (IXI) dataset, the Cambridge Centre for Ageing and Neuroscience (CamCAN), the Southwest University Adult Lifespan Dataset (SALD), the Strategic Research Program for Brain Sciences (SRPBS) Multi-disorder MRI Dataset, the Autism Brain Imaging Data Exchange (ABIDE), the ADHD-200 Consortium, the Nathan Kline Institute–Rockland Sample (NKI-RS), the Open Access Series of Imaging Studies (OASIS), and curated OpenNeuro cohorts. We acknowledge the investigators, data contributors, and funding bodies that supported the collection, curation, and public sharing of these datasets. These data were used in accordance with their respective data-use terms, citation requirements, and community data-sharing practices. Data used in the external validation analyses were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Data collection and sharing for ADNI are 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. The investigators within ADNI contributed to the design and implementation of ADNI and/or provided data, but did not participate in the analysis or writing of this report. Data used in the external validation analyses were also obtained from the Parkinson's Progression Markers Initiative / Parkinson's Precision Medicine Initiative (PPMI) database. PPMI is sponsored and partially funded by The Michael J. Fox Foundation for Parkinson's Research, with support from Aligning Science Across Parkinson's and additional funding partners. The PPMI investigators and contributors participated in the design, implementation, and data collection of PPMI, but did not participate in the analysis or writing of this report. We further acknowledge the developers and maintainers of FreeSurfer and related open-source neuroimaging resources that supported image processing, hippocampal subfield segmentation, and quantitative neuroanatomical analysis in this study.

Author contributions

Authors sharing identical initials are identified by their full names in parentheses. Study design: X.D. (Xiaozhou Du), L.C., Z.Z. (Zhizheng Zhuo), X.-N.Z. and Y.L. (Yaou Liu). Data analysis: X.D. (Xiaozhou Du), Z.Z. (Zhizheng Zhuo), Y.W. (Yinshan Wang), P.G., J.W. (Jinyuan Weng), R.Z. and Y.D. (Yunyun Duan). Statistical analysis: X.D. (Xiaozhou Du), Z.Z. (Zhizheng Zhuo), L.C., Y.W. (Yinshan Wang), and P.G. Data collection: X.D. (Xiaozhou Du), P.G., L.C., Z.Z. (Zhiqiang Zhang), X.Z. (Xinxiang Zhao), W.W., P.H., Y.L. (Yuerong Lizhu), Y.W. (Yinshan Wang), L.A., F.A., Y.B. (Yalin Bai), Y.B. (Yutong Bai), H.B., Q.C., J.C. (Jiancang Cao), F.C. (Feiyan Chen), F.C. (Feng Chen), K.C., Y.C., D.C., Z.C., H.D., Y.D. (Yuhui Du), G.F., Y.F., Z.G., L.G., C.G., M.G., Y.G., X.H., H.H., Y.H., B.H., J.H., C.-C.H., J.L. (Junqiang Lei), H.-J.L., J.L. (Jianrui Li), J.L. (Junlin Li), K.L. (Kexuan Li), S.L. (Shaowu Li), W.L. (Wenling Li), X.L., Y.L. (Yang Li), Y.L. (Yao Li), Y.L. (Yongmei Li), Y.L. (Yunfei Li), P.L., H.L., C.-P.L., J.L. (Jun Liu), J.L. (Jungang Liu), W.L. (Weiqi Liu), Y.L. (Ying Liu), S.L. (Shanshan Lu), J.L. (Jie Luo), Y.L. (Yuqi Luo), S.L. (Su Lui), H.M., N.M., L.M. (Lanxi Meng), L.M. (Linghui Meng), Y.Q., J.Q., J.S., Q.S., Y.S., Z.S., D.T., Y.T., F.W. (Fei Wang), J.W. (Jinhui Wang), X.X. (Xiaolu Xu), J.W. (Junkai Wang), X.W., Y.W. (Yilong Wang), Y.W. (Yong Wang), Z.W., J.W. (Jinyuan Weng), F.W. (Feiyun Wu), Y.W. (Ying Wu), Y.W. (Yuankui Wu), H.W., X.X. (Xiaona Xia), C.X., J.X. (Jiaheng Xu), J.X. (Jun Xu), R.X., S.X., T.G., R.Y. (Ruimeng Yang), S.Y., X.Y., Y.Y., Z.Y., F.Z. (Feifei Zhang), H.Z., J.Z. (Jianguo Zhang), N.Z., R.Z., T.Z., Y.Z., Z.Z. (Zhanjun Zhang), Z.Z. (Zhihua Zhang), J.Z. (Jian Zhao), L.Z. (Lianping Zhao), X.Z. (Xiaohu Zhao), X.Z. (Xin Zhao), X.Z. (Xia Zhe), J.Z. (Jiajia Zhu), L.Z. (Li Zhu), K.L. (Kuncheng Li), X.D. (Xujun Duan), Y.D. (Yunyun Duan), D.D., R.Y. (Rongjun Yu), F.Z. (Fuqing Zhou), Z.Z. (Zhizheng Zhuo), X.-N.Z. and Y.L. (Yaou Liu). Manuscript writing: X.D. (Xiaozhou Du). Manuscript editing: Z.Z. (Zhizheng Zhuo), L.C., Y.W. (Yinshan Wang), P.G., P.W., R.Y. (Rongjun Yu), J.C. (James Cole), S.H., X.-N.Z. and Y.L. (Yaou Liu).

Funding

This work was supported by a grant from the Beijing Outstanding Young Scientist Program (JWZQ20240101025) (Y.L.), the Chinese Institutes for Medical Research, Beijing (NO.CX23YZ01) (Y.L.), Healthy Longevity Catalyst Awards (HLCA/M-205/24) (R.Y.), the STI 2030-the major projects of the Brain Science and Brain-Inspired Intelligence Technology (2021ZD0200500) (X.N.Z.), the Key Program of the National Natural Science Foundation of China (NO. 82330057) (Y.L.), The Cross-Research Project of Beijing Science and Technology Star Program (NO. 20230484428) (Y.L.), the National Science Foundation of China (NO. 82202084 and NO. 82572171) (Z.Z.), Beijing Municipal Natural Science Foundation for Distinguished Young Scholars (No. JQ20035) (Y.L.), Ministry of Science and Technology of China-National Key Research Project (2022YFC2009904) (Y.L.), Beijing Hospital Management Center-Climb Plan (No. DFL20220503) (Y.L.), Beijing Young Scholars (Y.L.), Capital Medical University Young Scholars (Y.L.), Young Scientists Program of Beijing Tiantan Hospital, Capital Medical University (YSP202205) (Z.Z.), the Major Fund for International Collaboration of National Natural Science Foundation of China (81220108014) (X.N.Z.) and the National Basic Science Data Center 'Interdisciplinary Brain Database for In vivo Population Imaging' (ID-BRAIN) (X.N.Z.).

Competing interests

Jinyuan Weng is an employee of Philips Healthcare, Shanghai, China. Renlong Zhang is an employee of Neurosoft, Beijing, China. Weiqi Liu is an employee of Sophmind Technology (Beijing) Co Ltd, Beijing, China. The other authors declare that they have no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-79269-z.

Correspondence and requests for materials should be addressed to Demao Deng, Rongjun Yu, Fuqing Zhou, Zhizheng Zhuo, Xi-Nian Zuo or Yaou Liu.

Peer review information *Nature Communications* thanks Didac Vidal-Pineiro and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.

Xiaozhou Du[1,94], Peng Gao[2,94], Li Chai[1,94], Zhiqiang Zhang[3,94], Xinxiang Zhao[4,94], Wen Wang[5,94], Peiyu Huang[6,94], Yuerong Lizhu[1,94], Ping Wang[2], Yinshan Wang[2], Liping Ai[4], Feng Ao[7], Yalin Bai[8], Yutong Bai[9], Haihua Bao[10], Qing Cai[11], Jiancang Cao[12], Feiyan Chen[13], Feng Chen[14], Kun Chen[15], Yaojing Chen[2], Dan Cheng[16], Zaixu Cui[17], Hui Dai[18], Yuhui Du[19], Guoguang Fan[20], Ying Fu[21,22], Zheqi Gao[8], Lubin Gou[23,24], Chunjie Guo[25], Min Guo[1], Ya Guo[10], Xuemei Han[26], Hongjian He[13], Yang Hou[27], Basheng Hu[28], Jianping Hu[29], Chu-Chung Huang[11], Junqiang Lei[23,24], Hui-Jie Li[30], Jianrui Li[3], Junlin Li[31], Kexuan Li[32], Shaowu Li[33], Wenling Li[34], Xinyu Li[35], Yang Li[17], Yao Li[36], Yongmei Li[37], Yunfei Li[38], Peipeng Liang[39], Haiyan Liao[40], Ching-Po Lin[41,42], Jun Liu[43,44], Jungang Liu[45], Weiqi Liu[46], Ying Liu[47], Shanshan Lu[48], Jie Luo[36], Yuqi Luo[1], Su Lui[49], Hailin Ma[50], Ning Mao[51], Lanxi Meng[33], Linghui Meng[52], Yinfeng Qian[53], Jun Qiu[1,47], Juwei Shao[54], Qin Shen[40], Yulu Shi[55], Zhenyu Shu[32], Decai Tian[55], Yiheng Tu[30,56], Fei Wang[14], Jinhui Wang[57,58,59], Xiaolu Xu[1], Junkai Wang[60], Xiaochun Wang[61], Yilong Wang[55], Yong Wang[62], Zheng Wang[63,64], Jinyuan Weng[65], Feiyun Wu[48], Ying Wu[66], Yuankui Wu[28], Haiyan Wu[15], Xiaona Xia[67], Cheng Xu[35], Jiaheng Xu[68], Jun Xu[55], Rui Xu[69], Siyao Xu[1], Tianyu Gao[1], Ruimeng Yang[70,71], Songtao Yin[31], Xi Yin[72], Yulong Yin[73], Zhen Yuan[15,74], Feifei Zhang[61], Hailian Zhang[75], Jianguo Zhang[9], Ningnannan Zhang[76], Renlong Zhang[77], Tong Zhang[68], Yaqin Zhang[78], Zhanjun Zhang[2], Zhihua Zhang[1,79], Jian Zhao[52], Lianping Zhao[12], Xiaohu Zhao[38], Xin Zhao[73], Xia Zhe[80], Jiajia Zhu[53], Li Zhu[34], Kuncheng Li[81], Xujun Duan[82], Yunyun Duan[1,83], Sven Haller[1,84,85,86], James Cole[87,88], Demao Deng[89,95] ✉, Rongjun Yu[90,91,95] ✉, Fuqing Zhou[92,95] ✉, Zhizheng Zhuo[1,83,95] ✉, Xi-Nian Zuo[2,93,95] ✉ & Yaou Liu[1,83,95] ✉

[1]Department of Radiology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China. [2]State Key Laboratory of Cognitive Neuroscience and Learning, Faculty of Psychology, Beijing Normal University, Beijing, China. [3]Department of Diagnostic Radiology, Affiliated Jinling Hospital, Medical School of Nanjing University, Nanjing, China. [4]Department of Radiology, The Second Affiliated Hospital of Kunming Medical University, Kunming, China. [5]Department of Radiology & Functional and Molecular Imaging Key Lab of Shaanxi Province, Fourth Military Medical University, Xi'an, China. [6]Department of Radiology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China. [7]Department of Radiology, Renmin Hospital, Hubei University of Medicine, Shiyan, China. [8]Department of Radiology, Yuci Renmin Hospital, Jinzhong, China. [9]Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China. [10]Department of Medical Imaging Center, Qinghai University Affiliated Hospital, Xining, China. [11]School of Psychology and Cognitive Science, East China Normal University, Shanghai, China. [12]Department of Radiology, Gansu Provincial Hospital, Lanzhou, China. [13]School of Physics, Zhejiang University, Hangzhou, China. [14]Department of Radiology, Hainan General Hospital, Haikou, China. [15]Centre for Cognitive and Brain Sciences and Department of Psychology, University of Macau, Macau SAR, China. [16]Department of Radiology, University of Washington School of Medicine, Seattle, WA, USA. [17]Chinese Institute for Brain Research, Beijing, China. [18]Department of Radiology, Affiliated Hospital of Zunyi Medical University, Zunyi, China. [19]School of Computer and Information Technology, Shanxi University, Taiyuan, China. [20]Department of Radiology, The First Hospital of China Medical University, Shenyang, China. [21]Department of Neurology and Institute of Neurology, First Affiliated Hospital of Fujian Medical University, Fuzhou, China. [22]Institute of Neuroscience and Fujian Key Laboratory of Molecular Neurology, Fujian Medical University, Fuzhou, China. [23]Department of Radiology, Lanzhou University First Affiliated Hospital, Lanzhou, China. [24]Gansu Province Clinical Research Center for Radiology Imaging, Lanzhou, China. [25]Department of Radiology, The First Hospital of Jilin University, Changchun, China. [26]Department of Neurology, China-Japan Union Hospital of Jilin University, Changchun, China. [27]Department of Radiology, Shengjing Hospital of China Medical University, Shenyang, China. [28]Department of Radiology, Nanfang Hospital, Southern Medical University, Guangzhou, China. [29]Department of Radiology, First Affiliated Hospital of Fujian Medical University, Fuzhou, China. [30]State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China. [31]Department of Imaging Medicine, Inner Mongolia Autonomous Region People's Hospital, Hohhot, China. [32]Department of Radiology, Zhejiang Province People's Hospital, Hangzhou, China. [33]Beijing Neurosurgical Institute, Beijing, China. [34]Department of Radiology, General Hospital of Ningxia Medical University, Yinchuan, China. [35]Department of Imaging, Shanxi Provincial People's Hospital, Taiyuan, China. [36]National Engineering Research Center of Advanced

Magnetic Resonance Technologies for Diagnosis and Therapy, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China. [37]Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China. [38]Department of Radiology, The Fifth People's Hospital of Shanghai Fudan University, Shanghai, China. [39]School of Psychology, Capital Normal University, Beijing, China. [40]Department of Radiology, The Second Xiangya Hospital, Central South University, Changsha, China. [41]Institute of Neuroscience, National Yang Ming Chiao Tung University, Taipei, Taiwan. [42]Department of Education and Research, Taipei City Hospital, Taipei, Taiwan. [43]Department of Radiology, Chongqing University Central Hospital, Chongqing, China. [44]Department of Radiology, Chongqing Emergency Medical Center, Chongqing, China. [45]Department of Radiology, Xiamen Children's Hospital (Children's Hospital of Fudan University at Xiamen), Xiamen, China. [46]Sophmind Technology (Beijing) Co. Ltd., Beijing, China. [47]Department of Radiology, The First Affiliated Hospital of USTC, University of Science and Technology of China, Hefei, China. [48]Department of Radiology, The First Affiliated Hospital with Nanjing Medical University, Nanjing, China. [49]Huaxi MR Research Center (HMRRC), Department of Radiology, West China Hospital of Sichuan University, Chengdu, China. [50]Tibet Autonomous Region Key Laboratory for High Altitude Brain Science and Environmental Acclimatization, Lhasa, China. [51]Department of Radiology, Yantai Yuhuangding Hospital, Qingdao University, Yantai, China. [52]Department of Radiology, Hebei Medical University Third Hospital, Shijiazhuang, China. [53]Department of Radiology, The First Affiliated Hospital of Anhui Medical University, Hefei, China. [54]The Affiliated Hospital of Yunnan University, Kunming, China. [55]Department of Neurology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China. [56]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China. [57]Institute for Brain Research and Rehabilitation, South China Normal University, Guangzhou, China. [58]Key Laboratory of Brain, Cognition and Education Sciences (Ministry of Education), Guangzhou, China. [59]Center for Studies of Psychological Application, South China Normal University, Guangzhou, China. [60]Department of Radiology, Aerospace Center Hospital, Beijing, China. [61]Department of Radiology, First Hospital of Shanxi Medical University, Taiyuan, China. [62]Department of Radiology, The First Hospital of Hebei Medical University, Shijiazhuang, China. [63]School of Psychological and Cognitive Sciences, Peking University, Beijing, China. [64]School of Biomedical Engineering, Hainan University, Sanya, China. [65]Philips (China) Investment Co. Ltd., Shanghai, China. [66]School of Medical Imaging, Shanxi Medical University, Taiyuan, China. [67]Department of Radiology, Qilu Hospital (Qingdao), Cheeloo College of Medicine, Shandong University, Qingdao, China. [68]Department of Radiology, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, China. [69]Department of Radiology, First Affiliated Hospital of Xinjiang Medical University, Urumqi, China. [70]Department of Radiology, The Second Affiliated Hospital, South China University of Technology, Guangzhou, China. [71]Department of Radiology, Guangzhou First People's Hospital, Guangzhou, China. [72]Medical Imaging Center, The First Affiliated Hospital of Shihezi University, Shihezi, China. [73]School of Psychology, Northwest Normal University, Lanzhou, China. [74]Faculty of Health Sciences, University of Macau, Macau SAR, China. [75]Department of Radiology, The Fifth People's Hospital of Qinghai Province, Xining, China. [76]Department of Radiology and Tianjin Key Laboratory of Functional Imaging, Tianjin Medical University General Hospital, Tianjin, China. [77]Institute of Research and Clinical Innovations, Neusoft Medical Systems Co. Ltd., Beijing, China. [78]Department of Radiology, the Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, China. [79]Department of Neurosurgery, Tsinghua University Yuquan Hospital (Tsinghua University Hospital of Integrated Traditional Chinese and Western Medicine), Beijing, China. [80]Department of MRI, Shaanxi Provincial People's Hospital, Xi'an, China. [81]Department of Radiology, Xuanwu Hospital, Capital Medical University, Beijing, China. [82]The Clinical Hospital of Chengdu Brain Science Institute, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China. [83]China National Clinical Research Center for Neurological Diseases, Beijing, China. [84]Centre d'Imagerie Médicale de Cornavin, Geneva, Switzerland. [85]Department of Surgical Sciences, Research Area of Radiology and Nuclear Medicine, Uppsala University, Uppsala, Sweden. [86]Faculty of Medicine, University of Geneva, Geneva, Switzerland. [87]Centre for Medical Image Computing, University College London, London, UK. [88]Dementia Research Centre, Queen Square Institute of Neurology, University College London, London, UK. [89]Department of Radiology, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China. [90]Life Science Imaging Center, Hong Kong Baptist University, Hong Kong SAR, China. [91]Department of Education and Psychology, Academy of Wellness and Human Development, Faculty of Arts and Social Sciences, Hong Kong Baptist University, Hong Kong SAR, China. [92]Department of Radiology, The First Affiliated Hospital, Nanchang University, Nanchang, China. [93]Developmental Population Neuroscience Research Center, IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China. [94]These authors contributed equally: Xiaozhou Du, Peng Gao, Li Chai, Zhiqiang Zhang, Xinxiang Zhao, Wen Wang, Peiyu Huang, Yuerong Lizhu. [95]These authors jointly supervised this work: Demao Deng, Rongjun Yu, Fuqing Zhou, Zhizheng Zhuo, Xi-Nian Zuo, Yaou Liu. ✉e-mail: demaodeng@163.com; rongjunyu@hkbu.edu.hk; ndyfy02301@ncu.edu.cn; 18612161210@163.com; xinian.zuo@bnu.edu.cn; liuyaou@bjtth.org