

Cerebellar aging is spatially heterogeneous and supports cognitive resilience in later life

Received: 16 December 2024

Accepted: 31 March 2026

Published online: 10 June 2026

 Check for updates

Federico d'Oleire Uquillas¹✉, Esra Sefik¹, Jakob Seidlitz^{2,3,4},
Edan Daniel Hertz¹, Rafael Romero-Garcia⁵, Varun Warriar^{6,7},
Richard A. I. Bethlehem⁶, Aaron F. Alexander-Bloch^{2,3,4},
Jonathan D. Cohen¹, Samuel S.-H. Wang¹, Jorge Sepulcre^{8,9,10},
Patrizia Vannini^{9,11} & Jesse Gomez¹

The cerebellum contains most of the brain's neurons and supports many functions, yet how it changes with age remains unclear. Here we used three brain imaging studies spanning 47,000 adults and examined how different parts of the cerebellum age and their relation to cognition. We characterized cerebellar aging using volumetry and the T1-weighted/T2-weighted ratio, and corroborated these findings with quantitative magnetic resonance imaging in an independent sample. We show a spatially heterogeneous pattern of aging in which specific association and motor-related regions show steeper relationships with age than other lobules. Greater cerebellar volume was associated with higher cognitive scores with increasing age, suggesting that cerebellar structure may provide brain reserve that helps maintain function despite aging. In patients with Alzheimer's disease, cerebellar volume was linked to cognition in individuals with lower amyloid burden, especially in those carrying two copies of the *APOE-ε4* risk gene. This supports a threshold-reserve model, in which the cerebellum helps sustain cognition until pathology becomes widespread. These results show that the cerebellum has an active role in healthy cognitive aging and resilience.

Global population aging underscores the need to identify neural mechanisms supporting cognitive longevity. The number of individuals older than 70 years is increasing faster than that of younger adults¹. Although research has focused largely on the cerebral cortex, converging evidence indicates that the cerebellum, long recognized for sensorimotor control, also supports higher-order cognitive and affective functions^{2,3}. Through closed-loop circuits with cortical association areas and sub-cortical nuclei⁴, the cerebellum is positioned to influence cognition

across the lifespan. Posterior cerebellar regions linked to prefrontal and parietal networks exhibit protracted maturation⁵, suggesting a potential role in late-life cognitive resilience.

The cerebellum is not spared by aging. Morphometric and postmortem studies show hemispheric and vermal shrinkage, and loss of Purkinje and granule cells^{6,7}; longitudinal imaging shows accelerating atrophy in advanced age⁸. Aging effects appear particularly pronounced in posterior regions^{9–11}, and cerebellar volume

¹Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA. ²Lifespan Brain Institute, Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA, USA. ³Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA. ⁴Department of Child and Adolescent Psychiatry and Behavioral Science, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. ⁵Department of Medical Physiology and Biophysics, Institute of Biomedicine of Seville (IBiS) HUVR/CSIC/University of Seville/CIBERSAM, ISCIII, Seville, Spain. ⁶Department of Psychology, University of Cambridge, Cambridge, UK. ⁷Department of Psychiatry, University of Cambridge, Cambridge, UK. ⁸Gordon Center for Medical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. ⁹Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. ¹⁰Department of Radiology and Biomedical Imaging, Yale School of Medicine, Yale University, New Haven, CT, USA. ¹¹Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. ✉e-mail: uquillas@princeton.edu

Table 1 | Sample demographics and imaging characteristics in the HCP-aging sample

	HCP-aging sample
<i>n</i>	708
Age (years)	60.05 (15.59)
Sex, <i>n</i> (% female)	400 (56.5)
Total cerebellar volume (mm ³)	130,746.01 (13,494.55)
Total neocortical volume (mm ³)	456,301.18 (51,798.86)
eTIV (mm ³)	1,520,809.62 (188,357.16)
MoCA total score	26.32 (2.52)
MoCA-visuospatial score	4.14 (0.98)
Years of education	17.48 (2.21)
Adjusted household income	63,681.88 (62,350.68)

Data are reported as the mean (s.d.) unless stated otherwise.

correlates with executive function and processing speed in older adults^{10,11}. Yet, whether cerebellar variation contributes to cognitive reserve remains unclear because most prior studies used modest samples or coarse anatomical resolution, limiting assessment of spatial heterogeneity.

Reserve frameworks emphasize neural efficiency and adaptability as buffers against decline¹², but focus largely on neocortical and hippocampal systems. The ‘dysmetria of thought’ hypothesis¹³ proposes that cerebellar dysfunction disrupts higher-order cognition analogously to motor dyscoordination, implying that degeneration may produce cognitive dysmetria whereas preserved cerebellar integrity may confer protection. The cerebellum has also been implicated in Alzheimer’s disease (AD)¹⁴, but while neuroimaging studies suggest cerebellar involvement in AD¹⁵, whether it contributes to pathology or resilience¹⁶ remains unclear. Some work even reports slower cerebellar than neocortical structural decay¹⁷, although those analyses treated the cerebellum as a single structure and may have overlooked regional heterogeneity. Reduced cerebellar volume across neurodegenerative disorders¹⁸ further underscores the need to clarify its role in aging.

In this study, we combine large-scale multimodal data from the Human Connectome Project (HCP) (HCP-Aging, *n* = 708), UK Biobank (UKB) (*n* = 45,013) and the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (*n* = 1,423) to model cerebellar aging and its relationship to cognition, amyloid burden and the *APOE* genotype. We asked: Is cerebellar aging spatially heterogeneous? Do individual differences in cerebellar aging relate to cognitive resilience that is consistent with reserve? How do these processes manifest in AD dementia? By integrating normative and pathological aging, we position the cerebellum within broader frameworks of cognitive aging and reserve.

Results

Heterogeneous aging of cerebellar tissue and associations with cognition

In 708 neurotypical adults aged 36 to 100 years (Table 1), we examined age-related changes in cerebellar volume relative to the neocortex, controlling for biological sex, estimated total intracranial volume (eTIV) and years of education. Age was inversely associated to neocortical volume ($\beta = -1,927.92$, $P = 0.00001$, Cohen’s $f = 1.27$) and cerebellar volume ($\beta = -309.13$, $P < 0.00001$, Cohen’s $f = 0.54$), but not with eTIV ($P = 0.399$) (Fig. 1a–d).

To assess spatial heterogeneity, we parcellated the cerebellum into 11 rostrocaudal lobules. A mixed-effects model of log-transformed lobular volume (covariates: sex, eTIV, years of education, hemisphere) showed strong age-related volume associations and significant heterogeneity across lobules (age $\times$ lobule:

$F(1,014,826) = 50.07$, $P < 2.2 \times 10^{-16}$) (Fig. 1e,f). Main effects were present for age ($F(1,702) = 147.31$, $P < 2.2 \times 10^{-16}$), lobule ($F(1,014,826) = 65,756.26$, $P < 2.2 \times 10^{-16}$), hemisphere ($F(114,826) = 64.04$, $P = 1.31 \times 10^{-15}$), sex ($F(1,702) = 7.48$, $P = 0.006$), eTIV ($F(1,702) = 327.02$, $P < 2.2 \times 10^{-16}$) and education ($F(1,702) = 12.31$, $P = 0.0005$). Pairwise slope comparisons were Benjamini–Hochberg-corrected (Supplementary Table 1).

Allowing hemisphere-specific slopes did not alter lobular heterogeneity (age $\times$ lobule: $F(1,014,805) = 51.39$, $P < 2.2 \times 10^{-16}$). There was no overall hemispheric difference in age slopes (age $\times$ hemisphere: $F(114,805) = 1.53$, $P = 0.216$) and no age $\times$ lobule $\times$ hemisphere interaction ($F(1,014,805) = 1.27$, $P = 0.239$) (Supplementary Fig. 1).

We next examined vermis regions using the same log-volume framework. Mixed-effects models revealed heterogeneous age slopes across vermis regions (age $\times$ vermis: $F(42,820) = 4.56$, $P = 0.001$), alongside a main effect of age ($F(1,702) = 17.20$, $P = 0.00004$) (Supplementary Fig. 2). eTIV ($F(1,702) = 161.74$, $P < 2.2 \times 10^{-16}$) and education ($F(1,702) = 5.48$, $P = 0.019$) were significant covariates, whereas sex was not ($P = 0.166$). Most vermis regions declined by approximately 0.9–1.5% per decade. Vermis VI, VII, IX and X showed significant negative estimates after Benjamini–Hochberg correction, whereas vermis VIII was near zero (Supplementary Table 2).

To assess whether volumetric reductions reflected loss of fine-scale tissue structure, we analyzed cerebellar T1-weighted (T1W)/T2-weighted (T2W) ratio maps, a contrast sensitive to myelin and neurite density^{19,20}. Although T1W/T2W remains susceptible to cerebrospinal fluid partial-volume effects, it reduces intensity inhomogeneity and enhances tissue contrast. T1W/T2W declined with age across lobules (Fig. 1g): mixed-effects models (covariates: sex, eTIV, education and hemisphere) showed heterogeneous age slopes (age $\times$ lobule: $F(1,014,826) = 53.75$, $P < 2.2 \times 10^{-16}$), with main effects for lobule ($F(1,014,826) = 2,819.43$, $P < 2.2 \times 10^{-16}$), hemisphere ($F(114,826) = 1,860.92$, $P < 2.2 \times 10^{-16}$), and education ($F(1,702) = 10.79$, $P = 0.001$). Sex and eTIV were not significant (both $P > 0.8$). Percentage change per decade (Fig. 1h and Supplementary Table 3) showed significant decline across all regions after Benjamini–Hochberg correction, with the largest decreases in crus I (3.13%), crus II (2.98%) and lobule IX (3.14%), and the smallest in the anterior motor regions, consistent with a cerebellar sensory-association gradient⁵.

We next compared cerebellar aging to neocortical organization along the sensory-association hierarchy^{21,22}. For each cortical region, we related mean T1W/T2W in 360 young adults (a proxy of sensory-association rank) to its age-related T1W/T2W slope in the aging cohort (Fig. 1i). Regions with lower baseline T1W/T2W—typically association cortices—showed more negative age slopes, whereas highly myelinated sensory regions were relatively preserved.

Because T1W/T2W is a relative metric, we validated these patterns using quantitative magnetic resonance imaging (qMRI) in an independent cohort (*n* = 23; 29–71 years). We derived maps of the volume of interacting protons (VIP), sensitive to tissue water macromolecular interactions expected to decline with tissue loss. Mixed-effects models controlling for sex and hemisphere showed significant lobular heterogeneity (age $\times$ lobule: $F(10,462) = 5.10$, $P = 4.4 \times 10^{-7}$) (Fig. 1j), with main effects of lobule ($F(10,462) = 48.31$, $P < 2 \times 10^{-16}$) and hemisphere ($F(1,462) = 4.97$, $P = 0.026$), but not age ($F(1,20) = 3.61$, $P = 0.072$). Simple slopes (Supplementary Table 4) showed reductions in crus I, crus II, VIIb, VIIIa, VIIIb and IX (uncorrected $P < 0.05$). After Benjamini–Hochberg correction, VIIb and VIIIa remained significant ($P < 0.05$), with crus I and II, VIIIb and IX trending ($P < 0.1$). A posterior (crus I and II) versus anterior (I–III, IV) contrast confirmed faster decline in posterior cognitive lobules (estimate = -0.002 per decade, $t = -4.30$, $P < 0.0001$), an approximately two percentage point greater decline per decade than the anterior regions. These qMRI findings replicate the posterior-biased pattern observed with T1W/T2W.

To clarify the biological interpretation of VIP, we examined age associations for tissue volume fraction (TV), water tissue fraction (WTF) and VIP across cerebellar lobules (Supplementary Fig. 9). TV and WTF exhibited complementary age-related patterns consistent with shifts in macromolecular versus water content across the adult lifespan. VIP demonstrated partially overlapping but distinct age associations and several lobules showed VIP age effects despite minimal TV or WTF slopes. These findings suggest that VIP captures microstructural variation beyond simple tissue loss or water accumulation, providing a complementary and partial-volume-robust index of cerebellar microstructural aging.

To assess whether cerebellar structure was related to cognition, older adults completed the Montreal Cognitive Assessment (MoCA), a battery sensitive to cognitive decline. Higher MoCA scores were associated with greater neocortical volume ($\beta = 0.00001$, $P < 2.1 \times 10^{-9}$, Cohen's $f = 0.23$) (Fig. 1k) and greater cerebellar volume ($\beta = 0.0001$, $P < 3.1 \times 10^{-10}$, Cohen's $f = 0.24$) (Fig. 1m). The eTIV measure was also positively associated with MoCA performance ($\beta = 0.002$, $P = 0.011$, Cohen's $f = 0.10$) (Fig. 1l). Thus, greater global tissue volume, cortical and cerebellar, was related to cognitive performance in later life (Fig. 1n).

Regional cerebellar age-related volume changes and associations with cognition

Relating lobular gray matter volume (ACAPULCO regions of interest (ROIs)) to MoCA scores revealed heterogeneous structure cognition associations across lobules (Fig. 2a). In a joint log-volume model including all 11 bilateral lobules (adjusting for sex, eTIV, education and hemisphere), the MoCA $\times$ lobule interaction was significant ($F(10,692) = 2.19$, $P = 0.0169$), indicating spatially varying associations (Fig. 2b). Slopes are expressed as change in MoCA per 10% increase in lobular volume ($\Delta\text{MoCA}/10\%$), with Benjamini–Hochberg-adjusted P values. Lobules V and VI and crus I and II showed the strongest effects (VI: $+0.27$ (0.19, 0.35, 95% CI), $P = 5.9 \times 10^{-10}$; crus I: $+0.27$ (0.19, 0.35 CI), $P = 3.0 \times 10^{-10}$; crus II: $+0.21$ (0.14, 0.28 CI), $P = 4.4 \times 10^{-8}$) (Fig. 2b and Supplementary Table 5). An omnibus equality-of-slopes test confirmed

heterogeneity ($F(10,121.94) = 2.19$, $P = 0.017$). A priori contrasts demonstrated stronger associations in posterior/association lobules (crus I, crus II) relative to anterior/motor lobules (I–III, IV) (difference in mean slope = $+0.26$ MoCA per 10% (95% CI 0.12, 0.41), $z = 3.57$, $P = 0.0004$). Thus, a 10% larger posterior lobule corresponded to roughly 0.2–0.3 higher MoCA points.

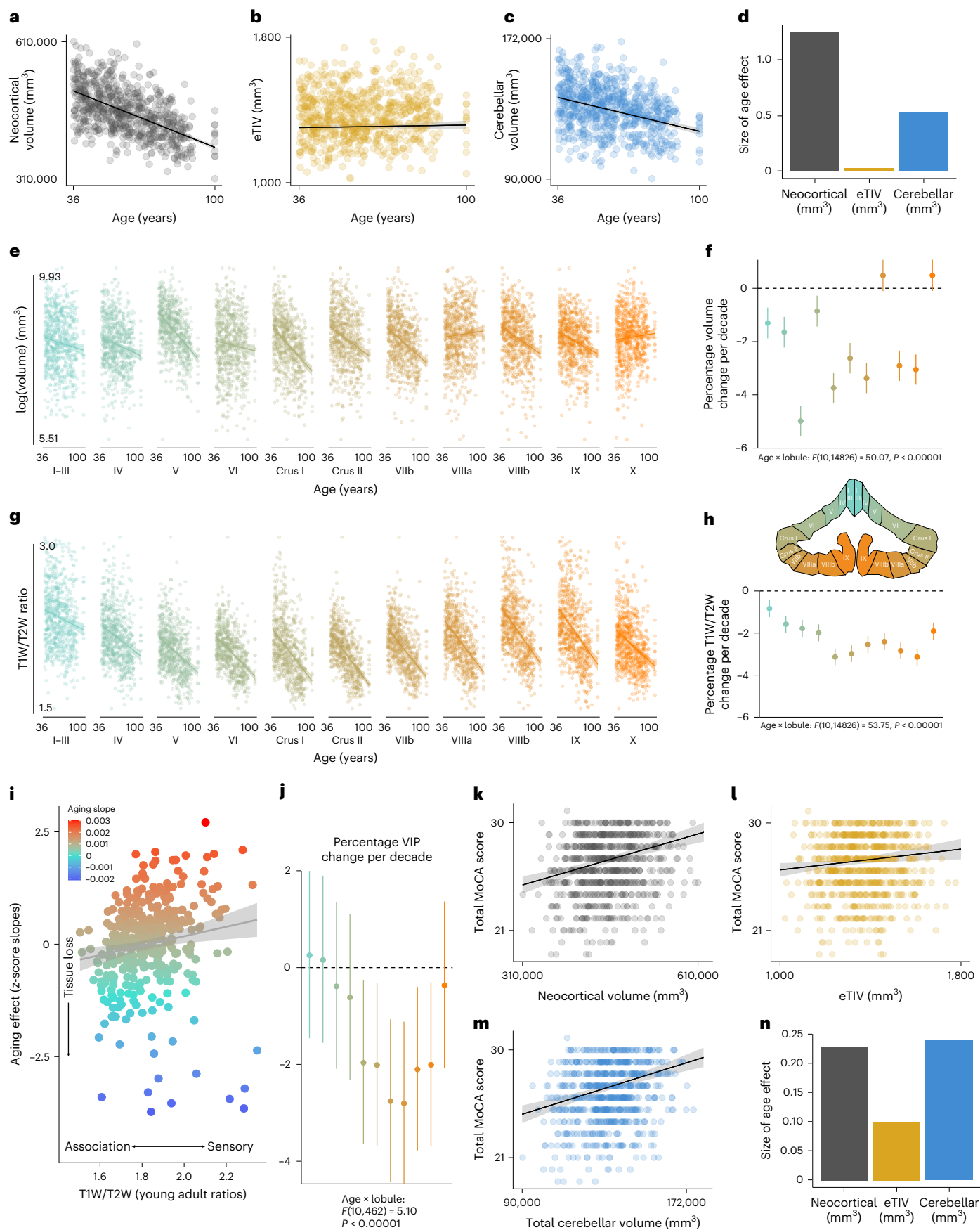
Voxel-based morphometry (VBM) corroborated these findings. Age-related gray matter associations were evident in the posterior cerebellum (family-wise error (FWE)-corrected $P < 0.0001$, false discovery rate-adjusted q value (qFDR) = 0.002, $F(1,705)$, cluster extent $k = 21$ voxels) (Fig. 2c, top left); MoCA was associated with bilateral lobules V, VI, crus I and crus II (FWE-corrected $P < 0.0001$, qFDR = 0.002, $F(1,705)$, $k = 21$) (Fig. 2c, bottom left). A conjunction analysis identified cerebellar cortex showing both aging and cognitive effects (Fig. 2c, right). Subscale analyses (Supplementary Fig. 3; unthresholded maps showing positive associations) revealed shared and distinct patterns. Only the visuospatial subscale survived multiple-comparisons correction (FWE-corrected $P < 0.0001$, qFDR = 0.002).

Network overlays showed conjunction clusters predominantly within association regions (frontoparietal, ventral-attention and portions of default mode) in the Buckner 7 Network parcellation²³, and within attention and language representations as per Multimodal Domain Task Battery parcellation², with some involvement of primary motor zones (Fig. 2d,e). Within these posterior regions, gray matter volume related to better cognition for both total MoCA ($\beta = 8.61$, $P = 1.4 \times 10^{-10}$, Cohen's $f = 0.25$) and the visuospatial subscale ($\beta = 3.86$, $P = 3.1 \times 10^{-16}$, Cohen's $f = 0.32$) (Fig. 2f,g).

MoCA scores declined linearly with age ($\beta = -0.04$, $P = 9.7 \times 10^{-12}$, Cohen's $f = 0.26$). Therefore, we tested whether cerebellar volume moderated age–cognition associations. Cerebellar volume attenuated age–cognition associations (age $\times$ cerebellar volume: total MoCA $P = 0.020$; visuospatial MoCA $P = 0.001$; Fig. 2f,g). Individuals with cerebellar volume one standard deviation below the mean showed steeper negative age–cognition slopes, whereas those one standard deviation above the mean exhibited attenuated

Fig. 1 | Cerebellar lobular heterogeneity and neocortical loss track cognitive differences with age. **a,c**, Linear regression models (two-sided t -tests of regression coefficients) correcting for biological sex, years of education and eTIV revealed significant inverse relationships between age and neocortical volume ($\beta = -1,927.92$ mm³ per year ($-19,279$ mm³ per decade), $P = 0.00001$, Cohen's $f = 1.27$) (**a**), and cerebellar volume ($\beta = -309.13$ mm³ per year ($-3,090$ mm³ per decade), $P < 2 \times 10^{-16}$, Cohen's $f = 0.54$) (**c**). **b**, Age and eTIV, correcting for sex and years of education, were not significantly related ($P = 0.399$). Solid lines represent ordinary least-squares regression fits (model-estimated mean response); shaded bands represent the 95% confidence intervals (CIs) of the regression estimate. All tests were two-sided. **d**, The magnitude of the age effects is summarized as Cohen's f multivariate effect sizes. **e**, Aging-related effects on cerebellar lobular volume demonstrate heterogeneous age associations across lobules. The lines represent OLS regression fits (model-estimated mean volume), and the shaded bands denote the 95% CIs of the regression estimate. **f**, Lobule-specific age slopes are expressed as percentage volume change per decade derived from a linear mixed-effects model of log-transformed lobular volume. Analyses were performed in $n = 707$ independent participants, with the individual participant as the unit of analysis. The age $\times$ lobule interaction was evaluated using a two-sided Wald F -test from the mixed-model analysis of variance (ANOVA) ($F(1,014,826) = 50.07$, $P < 2.2 \times 10^{-16}$). The two-sided t -tests and 95% CIs used Satterthwaite-approximated degrees of freedom. P values for lobule-wise slopes and pairwise slope differences were adjusted for multiple comparisons using the Benjamini–Hochberg FDR. **g**, A similar pattern characterized the T1W/T2W signal, with strong inverse relationships along the rostro–caudal axis suggesting lower relative myelination in posterior parcels. Lines represent OLS regression fits (model-estimated mean signal); shaded bands represent the 95% CIs of the regression estimate. **h**, The inset cartoon of a coronal plane cerebellum shows the lobule colormap used in the scatterplot. Points show the estimated percentage change per decade derived from a linear mixed-effects

model including an age $\times$ lobule interaction and the same covariates. Analyses were conducted in $n = 707$ independent participants. The age $\times$ lobule interaction was evaluated using a two-sided Wald F -test ($F(1,014,826) = 53.75$, $P < 2.2 \times 10^{-16}$). Lobule-specific slopes were estimated with two-sided t -tests and Satterthwaite-approximated degrees of freedom. Benjamini–Hochberg FDR correction was applied to multiple comparisons. Points represent model-estimated slopes. Whiskers denote the 95% CIs. **i**, Across 320 cortical ROIs, regions with lower young-adult T1W/T2W values (association-like cortex) showed more negative age-related slopes (greater tissue loss), whereas regions with higher T1W/T2W values (sensory-like cortex) showed flatter or positive slopes. Points represent ROI-specific regression slopes derived from two-sided general linear models. The line represents the model-estimated mean response; shaded bands denote the 95% CIs of the regression estimate. **j**, Lobule-specific age slopes in VIP are expressed as percentage volume change per decade, estimated from a linear mixed-effects model including an age $\times$ lobule interaction and controlling for sex and hemisphere. Analyses were conducted in $n = 23$ independent participants. The interaction was evaluated using a two-sided Wald F -test ($F(10,462) = 5.10$, $P = 0.0000004$). Lobule-specific slopes were estimated with two-sided t -tests and Satterthwaite-approximated degrees of freedom. Points represent model-estimated slopes. Whiskers denote the 95% CIs. **k,m**, Correcting for the same covariates as in **a**, greater MoCA total scores were associated with greater neocortical volume ($\beta = 0.01$ MoCA points per 1,000 mm³, $P = 2.13 \times 10^{-9}$, Cohen's $f = 0.23$) (**k**), and greater cerebellar volume ($\beta = 0.10$ MoCA points per 1,000 mm³, $P = 3.13 \times 10^{-10}$, Cohen's $f = 0.24$) (**m**). **l**, eTIV was also positively associated with cognition ($\beta = 0.20$ MoCA points per 100 cm³, $P = 0.011$, Cohen's $f = 0.10$). Effects were estimated using linear regression models (two-sided t -tests of regression coefficients). Points represent model-estimated regression coefficients; shaded bands denote the 95% CI of the regression estimate. **n**, Differences in predictive strength across neocortical, eTIV and cerebellar volume are displayed as Cohen's f effect sizes derived from the same regression models.



age–cognition slopes. This protective effect emerged after approximately 46 years of age for total MoCA and 44 years for visuospatial performance (simple slopes, positive FDR < 0.05). All moderation analyses were cross-sectional.

Cerebellar versus cortical contributions to cognitive reserve

We next tested whether the reserve effect was specific to the cerebellum or shared by neocortical regions functionally connected to posterior cerebellar lobules VI, crus I and crus II. Based on established cerebello-cerebral networks, we defined a posterior midline cerebral cortex composite (precuneus-postcingulate cortex (PPCC)). Cerebellar and PPCC volumes were strongly correlated ($\beta = 4,019.95$, $P = 2.0 \times 10^{-16}$, Cohen's $f = 0.41$), consistent with known subcortical-cortical networks²⁴. While age was negatively associated with PPCC volume ($\beta = -30.50$, $P = 2.0 \times 10^{-16}$, Cohen's $f = 0.89$) (Fig. 2h), and PPCC volume positively predicted MoCA ($\beta = 0.001$, $P = 3.7 \times 10^{-7}$, Cohen's $f = 0.19$) (Fig. 2i), PPCC volume did not moderate cognition–age associations (PPCC $\times$ age: $P = 0.263$) (Fig. 2j). Importantly, when PPCC volume was included as a covariate in the cerebellar moderation model, the age $\times$ cerebellar volume interaction remained significant ($P = 0.018$) (Fig. 2k), indicating that cerebellar reserve effects are not solely explained by structurally coupled cortical regions.

We extended this analysis to additional network-matched cortical composites: a frontoparietal network (FPN) composite (rostral and caudal middle frontal, inferior parietal cortex) and a ventral-attention network (VAN) composite (insula, inferior frontal gyrus pars opercularis/triangularis, supramarginal cortex). Neither FPN nor VAN volume moderated age–MoCA associations (FPN $\times$ age: $P = 0.188$; VAN $\times$ age: $P = 0.550$). In models including both cerebellar and cortical interaction terms, the age $\times$ cerebellar volume effect remained significant or marginal (PPCC model: $\beta = 0.014$, $P = 0.037$; FPN model: $\beta = 0.013$, $P = 0.056$; VAN model: $\beta = 0.016$, $P = 0.016$), whereas cortical interaction terms remained nonsignificant (PPCC: $P = 0.917$; FPN: $P = 0.972$; VAN: $P = 0.834$).

Fig. 2 | Cerebellar cognitive signature includes regions related to cerebello-cerebral networks and acts as a reserve factor. a, MoCA-related effects on ACAPULCO lobular volumes (mm^3) demonstrate a heterogeneous pattern in the association between MoCA scores and cerebellar volume. Models were fitted to log-transformed lobular volume; lines represent model-estimated mean responses from linear mixed-effects models. Shaded bands denote the 95% CIs of the regression estimate. All tests were two-sided. **b**, The points represent lobule-specific regression coefficients derived from a multiple linear regression predicting MoCA total score from log-transformed lobular volumes, correcting for sex, eTIV, years of education and hemisphere. Analyses were conducted in $n = 707$ independent participants, with the individual participant as the unit of analysis. Points show model-estimated coefficients expressed as the change in MoCA score per 10% increase in lobular volume (model fit used unscaled mm^3 predictors; coefficients scaled by $\log(1.10)$). The error bars denote the 95% CIs of the parameter estimate derived from two-sided t -tests of the regression coefficients. An omnibus Wald test of equality of lobule coefficients was performed (two-sided F -test), assessing whether associations between lobular volume and MoCA differed across lobules ($F(10,692) = 2.19$, $P = 0.0169$). Pairwise differences between lobule coefficients were evaluated using Wald tests based on the model covariance matrix, with P values adjusted for multiple comparisons using the Benjamini–Hochberg FDR. Lobules IV, V, VI, crus I, crus II, VIIb, VIIIb, IX and X, showed positive associations, whereas I–III and VIIIa were positive but their CIs overlapped zero. **c**, VBM using modulated, normalized gray matter maps, provided F -test association maps of clusters in the cerebellar gray matter (the color intensity reflects the magnitude of the F -statistic). Top left: F -test association map of clusters in the cerebellar gray matter associated with years of age, correcting for eTIV (FWE-corrected $P < 0.0001$, qFDR = 0.002, $F(1,705)$, $k = 21$). A brighter color (yellow) indicates greater years of age relating to lower cerebellar volume (mm^3). Bottom left: F -test map of clusters in the cerebellar gray matter associated with MoCA total scores, correcting for eTIV

We further tested canonical regions implicated in cognitive aging: the hippocampus and superior frontal cortex (Supplementary Fig. 4). Both showed expected negative age associations (hippocampus: $\beta = -14.91$, $P < 2.0 \times 10^{-16}$; superior frontal cortex: $\beta = -87.82$, $P < 2.0 \times 10^{-16}$), and positive associations with MoCA (hippocampus: $\beta = 0.001$, $P = 0.00001$; superior frontal cortex: $\beta = 0.0002$, $P = 0.00001$), yet neither moderated age–MoCA relationships (hippocampal volume $\times$ age: $P = 0.906$; superior frontal volume $\times$ age: $P = 0.831$). Critically, the age $\times$ cerebellar volume interaction remained significant after adjusting for hippocampal ($\beta = 0.163$, $P = 0.021$) or superior frontal volume ($\beta = 0.169$, $P = 0.017$). Together, these results demonstrate that cerebellar volume contributes unique protective variance to age-related cognition associations beyond structurally and functionally related neocortical regions.

Assessment of sex effects on MoCA performance, education, eTIV and volume

In models correcting for eTIV and years of education, male and female HCP participants did not differ in neocortical volume ($P = 0.990$; male cortex mean volume: $478,998.40 \text{ mm}^3$; female cortex mean volume: $438,824.30 \text{ mm}^3$), cerebellar volume ($P = 0.071$; male cerebellar mean volume: $137,083.60 \text{ mm}^3$; female cerebellar mean volume: $125,866.10 \text{ mm}^3$), education ($P = 0.501$; male mean years of education: 17.80; female mean years of education: 17.23) or age ($P = 0.980$; male mean age: 60.35; female mean age: 59.82). As expected, females exhibited smaller eTIV ($\beta = -73.83$, Cohen's $f = 0.67$, $P = 2.0 \times 10^{-16}$; male eTIV mean: $1,454.23 \text{ mm}^3$; female eTIV mean: $1,302.61 \text{ mm}^3$). Females also scored modestly higher on the MoCA ($\beta = 0.494$, Cohen's $f = 0.17$, $P = 0.00002$; female MoCA average: 26.55; male MoCA average: 26.01).

We next tested whether sex moderated imaging–cognition associations. Sex did not interact with neocortical volume ($P = 0.161$), cerebellar volume ($P = 0.912$), age ($P = 0.563$), education ($P = 0.093$) or eTIV ($P = 0.274$) (Supplementary Fig. 5). Cerebellar VBM gray matter age maps, corrected for eTIV, were similar across sexes (Supplementary Fig. 5g, top; FWE-corrected $P < 0.05$). Although

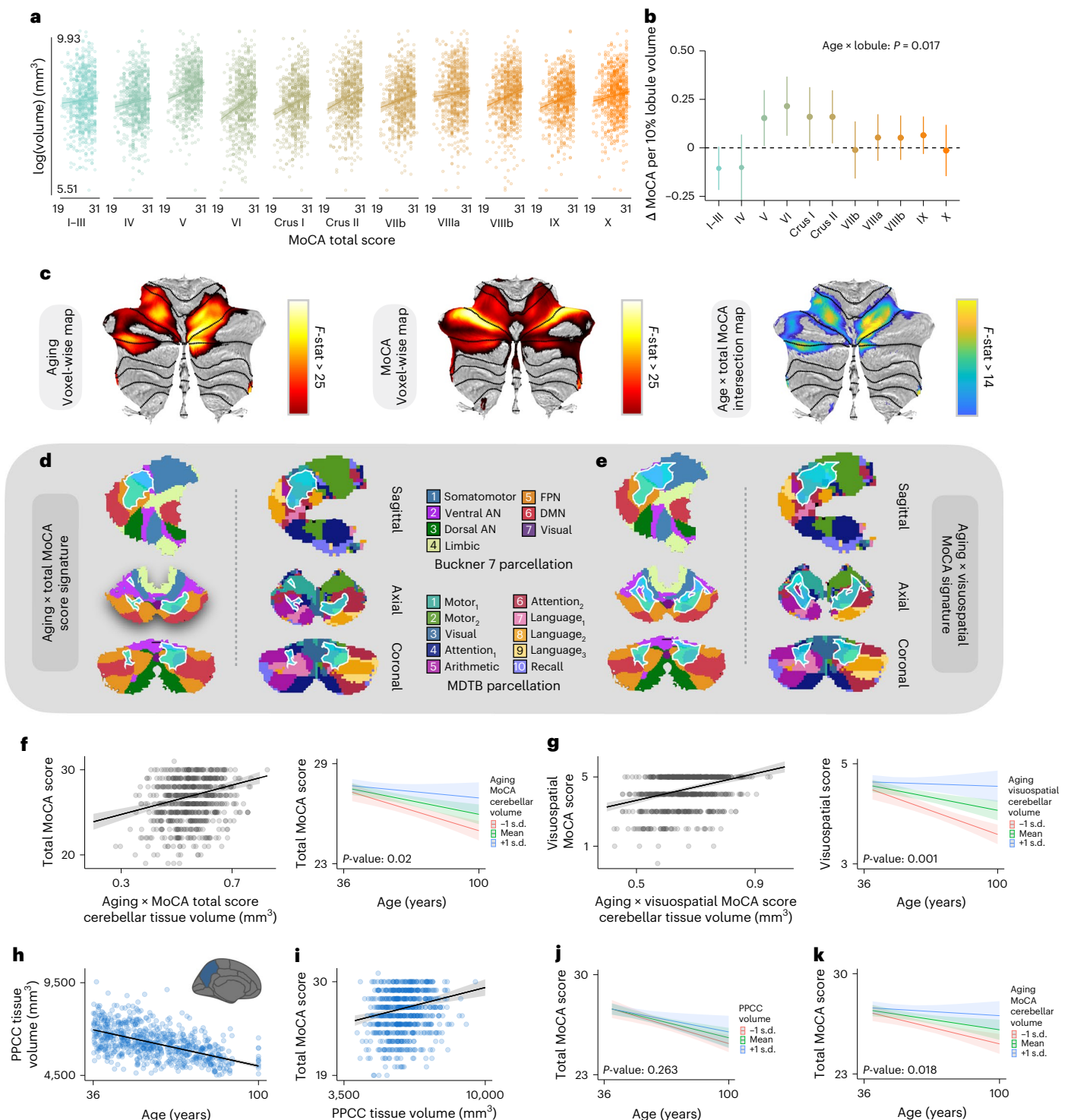
(FWE-corrected $P < 0.0001$, qFDR = 0.002, $F(1,705)$, $k = 21$). Right: conjunction map (FWE-corrected $P < 0.0001$, qFDR = 0.002, $F(1,705)$, $k = 21$) visualizing the intersection of the aging and MoCA voxel-wise results, suggests that variations in this map contribute to aging-related changes in cognitive performance as measured by the MoCA. **d,e**, Resulting conjunction intersection clusters (FWE-corrected $P < 0.0001$, qFDR = 0.002, $F(1,705)$, $k = 21$) of the aging–MoCA total score voxel-wise map (**d**) and the aging–MoCA-visuospatial subscale score map (**e**) overlaid on maps of Buckner 7 and Multimodal Domain Task Battery parcellations. The white outlines delineate the aging–MoCA mask in cyan. Overlaps illustrate qualitatively where clusters fall relative to broad functional zones. **f,g**, The lines represent model-estimated mean responses from the linear interaction models (two-sided tests of interaction terms); the shaded bands denote the 95% CIs of the regression estimate. **f**, Left: greater mean VBM-derived gray matter signal within the aging–total MoCA cerebellar signature mask was associated with higher total MoCA scores ($\beta = 8.61$ MoCA points per unit increase in signature value, $P = 1.35 \times 10^{-10}$, Cohen's $f = 0.25$). **g**, Left: greater mean VBM-derived gray matter signal within the aging–visuospatial cerebellar signature mask was associated with higher MoCA-visuospatial scores ($\beta = 3.86$ score units per unit increase in signature value, $P = 3.06 \times 10^{-16}$, Cohen's $f = 0.32$). **f**, Right: the negative association between age and MoCA scores was attenuated at higher levels of aging–MoCA signature ($P = 0.02$). **g**, Right: the aging–visuospatial cerebellar signature similarly moderated the association between age and visuospatial performance ($P = 0.001$). **h–k**, Post-hoc linear regression models, controlling for the same covariates, examined the relationship between PPCC volume, age, cognition and cerebellar reserve. **h**, Age was inversely related to PPCC volume ($\beta = -30.25 \text{ mm}^3$ per year, $P < 2 \times 10^{-16}$, Cohen's $f = 0.88$). **i**, Higher PPCC volume was associated with greater total MoCA scores ($\beta = 0.001$ MoCA points per mm^3 , $P = 3.71 \times 10^{-7}$, Cohen's $f = 0.19$). **j**, PPCC did not moderate the association between age and MoCA scores ($P = 0.263$). **k**, Aging–MoCA cerebellar signature remained a moderator of the age–MoCA association after including PPCC volume in the model ($P = 0.018$).

MoCA-related VBM maps showed a subtle right hemisphere bias, this effect did not survive correction for multiple comparisons (shown at uncorrected $P < 0.0001$; Supplementary Fig. 5g, bottom). Together, these findings indicate that the observed cerebellar aging and reserve effects are not driven by sex differences.

Generalizability of cerebellar reserve to independent cognitive tasks

To test whether the aging-MoCA cerebellar signature generalized to independent measures, we examined four HCP tasks: dimensional card sorting task²⁵; flanker attention task²⁶; picture sequence memory (PSM) task²⁷; and list sorting working memory (LSWM) task²⁸. Interaction

models controlling for sex, eTIV and education showed that greater cerebellar volume attenuated age-related decline in the card sorting (cerebellar volume $\times$ age: $\beta = 0.99$, $P = 0.0005$) and flanker task (cerebellar volume $\times$ age: $\beta = 0.710$, $P = 0.005$) (Supplementary Fig. 6a,b). Simple slopes (FDR < 0.05) indicated that protective effects emerged after approximately 55 (card sorting task) and 60 (flanker attention task) years (Supplementary Fig. 6b). Additive models showed positive main effects of cerebellar volume on both tasks (card sorting: $\beta = 32.84$, $P = 1.1 \times 10^{-9}$; flanker: $\beta = 31.29$, $P = 3.5 \times 10^{-10}$). By contrast, the aging-MoCA cerebellar signature volume did not moderate age-related performance for the PSM task ($P = 0.646$) (Supplementary Fig. 6c) or LSWM task ($P = 0.441$) (Supplementary Fig. 6d), although positive



cerebellar volume main effects were observed in additive models (PSM model main effect of cerebellar volume: $\beta = 45.81$, $P = 1.5 \times 10^{-7}$; LSWM model main effect of cerebellar volume: $\beta = 45.81$, $P = 1.5 \times 10^{-7}$).

Cerebellar cognitive reserve in the UKB

We replicated our findings in an independent, large dataset of healthy older adults from the UKB (Table 2). Focusing on posterior lobules (crus I, crus II, lobule IX), we observed substantial age-related volume decline, most pronounced in crus I ($\beta = -0.002$, $P < 2.0 \times 10^{-16}$, Cohen's $f = 0.38$), followed by crus II ($\beta = -0.002$, $P < 2.0 \times 10^{-16}$, Cohen's $f = 0.27$) and lobule IX ($\beta = -0.003$, $P < 2.0 \times 10^{-16}$, Cohen's $f = 0.16$) (Fig. 3a). In a joint mixed-effects model controlling for sex, hemisphere, education and total head volume, age effects differed across lobules (age $\times$ lobule: $F(2,181,270) = 309.34$, $P < 2.2 \times 10^{-16}$), with all main effects significant (all $P < 3.5 \times 10^{-13}$). Back-transformed estimates showed scaled-free decline across regions: crus I (-6.36% per decade, 95% CI -6.55 , -6.18), crus II (-4.68% , 95% CI -4.87 , -4.49) and lobule IX (-4.44% , 95% CI -4.63 , -4.25) (Fig. 3b), replicating strong aging effects in the posterior cerebellum.

Cognitive performance declined with age for both digit symbol substitution task (DSST)²⁹ ($\beta = -0.25$, $P < 2.0 \times 10^{-16}$, Cohen's $f = 0.41$) and the trail-making test part B (TMB)³⁰ ($\beta = 0.93$, $P < 2.0 \times 10^{-16}$, Cohen's $f = 0.31$) (Fig. 3c). Posterior lobules showed the strongest task correlations (crus I: DSST $r = 0.19$, TMB $r = -0.18$; crus II: DSST $r = 0.14$, TMB $r = -0.16$; IX: DSST $r = 0.13$, TMB $r = -0.12$) (Fig. 3d).

Greater volume attenuated age-related decline for both the DSST (age $\times$ cerebellar volume: $\beta = 0.01$, $P = 0.044$) (Fig. 3e) and the TMB (age $\times$ cerebellar volume: $\beta = -0.10$, $P = 0.00004$) (Fig. 3f). That is, at age 80, individuals with higher cerebellar volume completed the TMB approximately 9 s faster and performed about 5% better on the DSST. Thus, in a large independent cohort, analysis of cerebellar structure replicated the heterogeneous aging across lobules and the protective moderation of age-related cognitive decline observed in the HCP, now in a mega-dataset of over 40,000 adults.

Cerebellar reserve in AD

We next tested whether cerebellar reserve extends to clinical populations using the ADNI ($n = 1,423$) (Table 3). Participants were stratified according to amyloid status (amyloid-negative, $A\beta^-$; amyloid-positive, $A\beta^+$) and cerebellar volume $\times$ age interactions were examined for MoCA scores, controlling for sex and eTIV. In $A\beta^-$ individuals, larger cerebellar volume attenuated age-related MoCA slopes (cerebellar volume $\times$ age: $P = 0.0492$) (Fig. 4a). No moderation effect was detected in $A\beta^+$ participants ($P = 0.869$) (Fig. 4b), consistent with a threshold-reserve model in which cerebellar structure buffers cognition when amyloid burden is low.

MoCA scores declined across diagnostic stage (cognitive normal (CN), mild cognitive impairment (MCI), AD; $F(2) = 289.60$, $P < 0.0001$), with all pairwise differences significant (mean MoCA score in CN: 25.07; MCI: 22.36; AD: 16.99) (Supplementary Fig. 7a). Neocortical volume similarly decreased according to stage ($F(2) = 27.68$, $P < 0.0001$; mean cortex volume in CN: 427,996.10 mm³; MCI: 423,554.40 mm³; AD: 395,070.80 mm³) (Supplementary Fig. 7b). In contrast, cerebellar volume did not differ across diagnosis ($F(2) = 0.74$, $P = 0.475$; mean cerebellar volume in CN: 127,668.20 mm³; MCI: 128,135.30 mm³; AD: 126,777.90 mm³) (Supplementary Fig. 7c).

We next examined cerebellar volume–cognition associations stratified according to diagnosis and amyloid status. In $A\beta^-$ individuals, cerebellar volume interacted with diagnostic group in predicting MoCA ($A\beta^-$ AD group: $\beta = 0.0001$, $P = 0.004$) (Fig. 4c and Supplementary Table 6). However, in a three-way model (age $\times$ cerebellar volume $\times$ group), no significant age moderation was observed within subgroups (all $P > 0.3$).

Given that *APOE* $\epsilon 4$ is a major AD risk factor, we tested whether cerebellar reserve persisted across genotype. Among $A\beta^-$ individuals, greater cerebellar volume predicted higher MoCA scores across

Table 2 | Sample demographics and imaging characteristics in the UKB sample

	UKB sample
<i>n</i>	45,577
Age (years)	63.60 (7.55)
Sex, <i>n</i> (% female)	19,474 (52.9)
Cerebellar volume (mm ³)	6,963.05 (874.29)
Total brain volume (mm ³)	1,054,072.94 (128,909.05)
Years of education, <i>n</i> (%)	
Professional	20,552 (49.7)
NVQ, HND, HNC	5,760 (13.9)
CSEs	8,607 (20.8)
O Levels	1,854 (4.5)
A Levels	2,415 (5.8)
College	2,177 (5.3)
DSST correct matches score	20.24 (4.96)
TMB (s)	64.19 (23.31)
Mean household income, <i>n</i> (%)	
Less than 18,000	2,935 (9.9)
Less than 31,000	858 (2.9)
Less than 52,000	4,707 (15.9)
Less than 100,000	8,921 (30.1)
Greater than 100,000	12,261 (41.3)

Data are reported as mean (s.d.) unless stated otherwise. Sex (count/%) is reported per cohort; ADNI is presented according to diagnostic/amyloid subgroup. A Level, Advanced Level (13 years of study); CSE, Certificate of Secondary Education; HNC, Higher National Certificate; HND, Higher National Diploma; NVQ, National Vocational Qualification; O Level, Ordinary Level (11 years of study).

haplotypes, including $\epsilon 4/\epsilon 4$ homozygotes (Fig. 4d) ($\epsilon 4/\epsilon 4$: $\beta = 0.001$, $t = 3.14$, $P = 0.00175$; $\epsilon 3/\epsilon 3$: $\beta = 0.002$, $t = 3.11$, $P = 0.00193$; $\epsilon 3/\epsilon 4$: $\beta = 0.001$, $t = 3.09$, $P = 0.00210$). No significant effects were observed in $A\beta^+$ participants (all $P > 0.2$) (Fig. 4e), indicating that cerebellar reserve operates primarily in the relative absence of substantial amyloid burden, including in genetically at-risk individuals.

Supplementary mixed-effects models stratified according to $A\beta$ status (subject random intercepts, controlling for sex, eTIV and hemisphere), showed age $\times$ lobule interactions in both $A\beta^-$ ($F(1,014,763) = 15.35$, $P < 0.00001$) and $A\beta^+$ individuals ($F(1,013,545) = 14.08$, $P < 0.00001$). In both strata, posterior lobules (VI, VIIb, VIIIb, IX, crus I) showed steeper percentage-per-decade estimates than anterior lobules (I–IV), with the exception of strong effects in anterior lobule V (Supplementary Fig. 8a,b and Supplementary Table 7). Hemisphere did not moderate these effects (age $\times$ lobule $\times$ hemisphere: $A\beta^-$ $P = 0.220$; $A\beta^+$ $P = 0.516$).

Volume–MoCA associations were modest but spatially heterogeneous (volume $\times$ lobule: $A\beta^-$ $F(1,015,375) = 8.23$, $P < 0.00001$; $A\beta^+$ $F(1,014,143) = 7.19$, $P < 0.00001$) (Supplementary Table 8). In $A\beta^-$ participants, positive Δ MoCA per 10% larger volume was observed in lobule IV (+0.24%, $P = 0.003$), V (+0.39%, $P < 0.00001$), VI (+0.21%, $P = 0.008$), crus I (+0.28%, $P = 0.003$), crus II (+0.15%, $P = 0.029$), VIIb (+0.45%, $P < 0.00001$) and IX (+0.16%, $P = 0.003$) (Benjamini–Hochberg-adjusted $P < 0.05$). In $A\beta^+$ individuals, effects again concentrated in lobule V (+0.26%, $P = 0.0008$), crus I (+0.27%, $P = 0.003$) and VIIb (+0.42%, $P < 0.00001$), with most other regions null. Lobule X showed a small negative association (-0.25% , $P = 0.001$). Effects were modest (-0.2 – 0.45% MoCA per 10% larger volume).

Longitudinal ADNI analyses further showed that greater cerebellar volume predicted higher overall MoCA performance

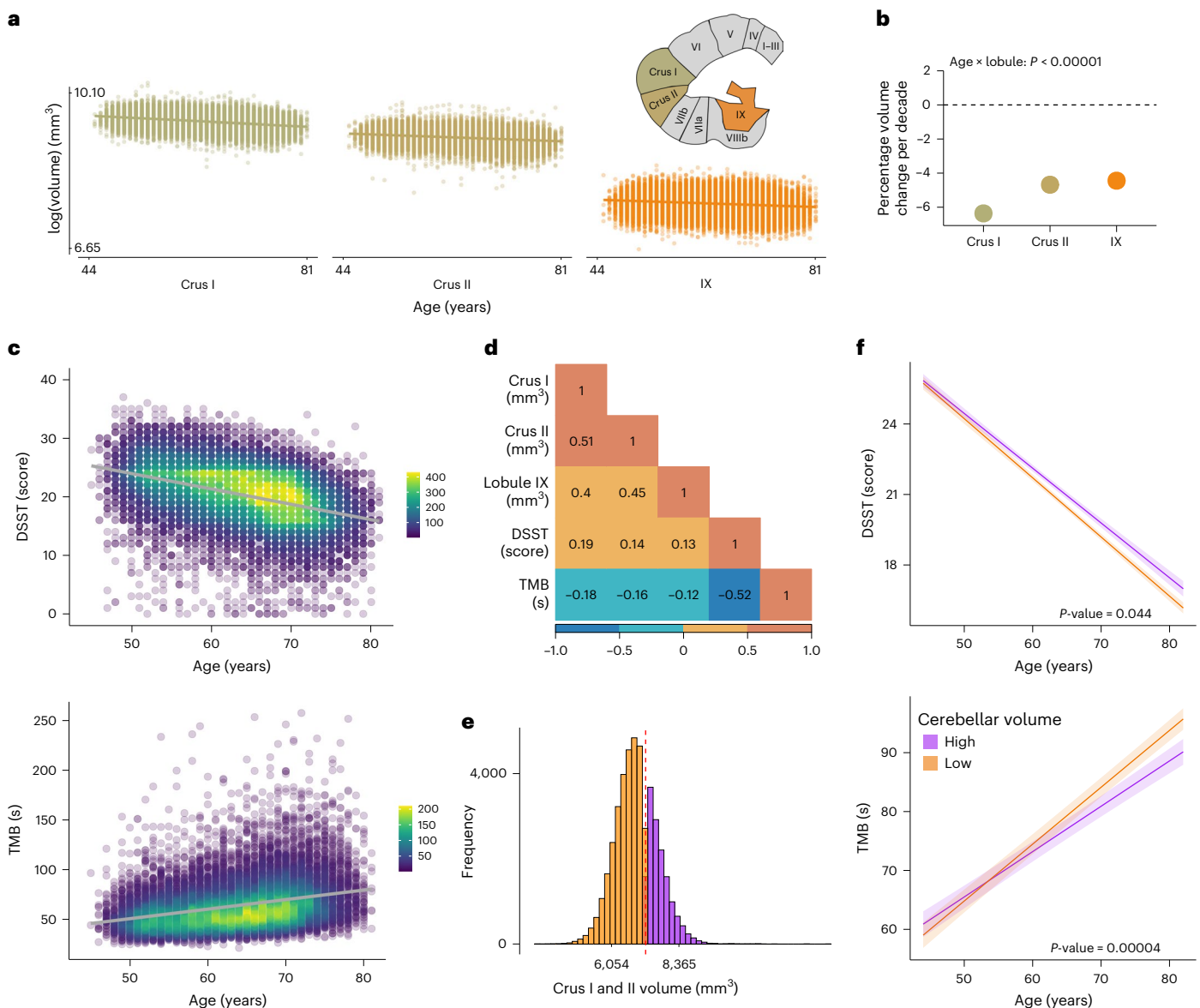


Fig. 3 | The cerebellum reproducibly modulates age-related cognitive differences in the UKB. Analyses were conducted in $n = 35,763$ independent participants with full complete data, with the individual participant as the unit of analysis. **a**, Replication of the HCP findings in the UKB cohort revealed heterogeneous age-related effects across crus I, crus II and lobule IX in linear models correcting for biological sex, total head volume and years of education. **b**, The points represent lobule-specific age slopes expressed as percentage volume change per decade, derived from a linear mixed-effects model of log-transformed lobular volume. The age × lobule interaction was evaluated using a two-sided Wald *F*-test from the mixed-model ANOVA (age × lobule: $F(2,181,270) = 309.34, P < 2.2 \times 10^{-16}$). Lobule-specific slopes were estimated with two-sided *t*-tests and the 95% CIs of the parameter estimate (smaller than the plotted data points). Estimated losses per decade were 6.36% for crus I (95% CI -6.36, -6.18), 4.68% for crus II (95% CI -4.87, -4.49) and 4.44% for lobule IX (95% CI -4.63, -4.25). **c**, Top: in linear regression models (two-sided *t*-tests),

greater age was associated with lower DSST scores ($\beta = -0.25$ DSST points per year, $P < 2 \times 10^{-16}$, Cohen's $f = 0.41$). **c**, Bottom: greater age was associated with longer completion time on the TMB ($\beta = 0.93$ seconds per year, $P < 2 \times 10^{-16}$, Cohen's $f = 0.31$). The lines represent model-estimated mean responses from linear regression models; the shaded bands denote the 95% CIs of the regression estimate. **d**, Unadjusted Pearson correlations demonstrated associations among lobular volumes (crus I, crus II and lobule IX) and between cognitive measures. **e**, A Gaussian mixture model applied to a crus I–II composite identified two distinct distributions, enabling classification into lower-volume and higher-volume groups. **f**, Linear interaction models (two-sided tests of interaction terms) correcting for sex, years of education and total head volume revealed that individuals with higher cerebellar volume exhibited lower DSST scores ($P = 0.044$) and longer completion time on the TMB ($P = 0.00004$). The lines represent model-estimated mean outcomes from the interaction model; the shaded bands denote the 95% CIs of the regression estimate.

($\beta = 0.23, P = 0.04$). Although volume × time was not significant in models with random intercepts and random slopes ($P = 0.421$), a significant interaction emerged in a random-intercept-only model ($\beta = -0.031, P = 0.019$), with higher initial cognitive performance and convergence over time. This pattern is consistent with a 'front-loaded' reserve model in which structural advantages buffer early cognitive differences.

Discussion

Across neurotypical aging and clinical populations, we show that the cerebellum exhibits spatially heterogeneous senescence, with posterior association lobules and lobule V disproportionately vulnerable to tissue loss. This spatial heterogeneity was replicated in an independent, large-scale dataset, indicating that it is a robust feature of subcortical aging. Although the pattern echoes the sensory-association hierarchy

Table 3 | Sample demographics and imaging characteristics in the ADNI sample

	CN AB ⁻	CN AB ⁺	MCI AB ⁻	MCI AB ⁺	AD AB ⁻	AD AB ⁺	P value
<i>n</i>	303	280	140	123	80	62	
Age (years)	74.50 (6.52)	75.09 (6.66)	77.01 (8.08)	75.59 (7.38)	75.59 (8.15)	75.60 (7.71)	0.028
Sex, <i>n</i> (% female)	168 (55.4)	157 (56.1)	56 (40.0)	45 (36.6)	30 (37.5)	28 (45.2)	<0.001
Total cerebellar volume (mm ³)	128,536.33 (12,534.48)	126,742.07 (11,900.30)	125,587.07 (12,507.90)	127,262.08 (13,347.47)	126,333.72 (11,792.44)	127,340.50 (11,955.39)	0.277
Total neocortical volume (mm ³)	429,978.70 (45,107.06)	425,880.94 (43,210.13)	403,448.18 (50,591.46)	410,926.35 (56,019.73)	389,682.71 (48,569.70)	401,895.65 (48,685.18)	<0.001
eTIV (mm ³)	1,451,505.59 (156,887.52)	1,438,760.03 (149,726.34)	1,465,296.78 (157,303.30)	1,482,354.46 (158,795.24)	1,422,662.74 (175,772.87)	1,474,327.26 (178,113.72)	0.060
Years of education	16.32 (2.78)	16.40 (2.65)	15.45 (2.89)	15.56 (3.44)	15.27 (2.66)	15.62 (2.75)	0.086
MoCA total score	24.94 (2.15)	25.21 (1.83)	20.21 (5.20)	20.52 (6.08)	16.87 (4.39)	17.13 (5.02)	<0.001
CDR							<0.001
0	263 (87.1)	241 (86.4)	18 (12.9)	11 (8.9)	0 (0)	1 (1.6)	
0.5	37 (12.3)	38 (13.6)	85 (60.7)	82 (66.7)	22 (27.5)	19 (30.6)	
1	2 (0.7)	0 (0)	28 (20.0)	20 (16.3)	52 (65.0)	30 (48.4)	
2	0 (0)	0 (0)	8 (5.7)	8 (6.5)	6 (7.5)	11 (17.7)	
3	0 (0)	0 (0)	1 (0.7)	2 (1.6)	0 (0)	1 (1.6)	

Data are reported as the mean (s.d.) unless stated otherwise.

of the neocortex, larger cerebellar volume attenuated age-related cognitive differences, consistent with a cerebellar contribution to brain reserve.

Spatial heterogeneity of cerebellar aging

Prior morphometric studies suggested regional variability in cerebellar aging^{9,10,31} but lacked systematic lobular comparisons and microstructural resolution. Across large, multi-cohort datasets and complementary tissue metrics, we identified a reproducible posterior-skewed pattern of aging, with association lobules (notably crus I and crus II) exhibiting steeper age-related effects than anterior motor lobules, with the notable exception of lobule V (Fig. 1e). This spatial pattern broadly mirrors functional cerebellar organization and provides a high-resolution, cross-cohort characterization of cerebellar aging.

The steep volumetric decline in lobule V, an anterior sensorimotor region, contrasted with stronger posterior microstructural effects in crus I and crus II, suggesting that volumetric and tissue-composition metrics capture partly distinct biological processes. Lobule V volume loss, which was observed in both the HCP and ADNI datasets (Supplementary Fig. 8), may reflect macroscopic degeneration, whereas posterior microstructural changes may reflect alterations in myelination or synaptic density. Intrinsic physiology may also contribute: Purkinje cells in the anterior lobules (III and IV) exhibit greater excitability than those in nodular regions such as lobule X³²; within lobule V, zebrin-positive and zebrin-negative modules differ markedly in firing rates, with zebrin-negative cells firing roughly twice as fast³³. Anterior lobules are also particularly susceptible to alcohol-related degeneration in adults³⁴, suggesting that lobule V may lie at a physiological intersection where elevated excitability and environmental vulnerability increase cumulative metabolic burden across the lifespan.

Microstructural analyses further indicated that age-related cerebellar volume loss reflects tissue-content changes rather than uniform shrinkage. Cerebellar lobules showed reductions in tissue indices with age, whereas neocortical sensory regions exhibited higher T1W/T2W values, plausibly reflecting iron accumulation³⁵. Cortical association areas instead displayed trajectories resembling posterior cerebellar lobules, consistent with coordinated vulnerability across distributed association systems. Preferential decline of later-maturing cerebellar regions is also consistent with retrogenesis^{36,37}, whereby

developmentally protracted systems are disproportionately susceptible to degeneration. Together, these complementary findings indicate that cerebellar aging is spatially organized, developmentally patterned and biologically grounded.

Cerebellar structure and cognitive performance

Cerebellar aging was behaviorally meaningful. Across cohorts, higher cognitive performance, including MoCA, processing speed and executive measures, was associated with greater cerebellar volume, particularly in posterior lobules representing frontoparietal, default-mode and ventral-attention networks¹⁰ (Fig. 2d). Cerebellar volume also moderated age–cognition associations, such that individuals with larger cerebella exhibited attenuated age-related cognitive differences (Fig. 2f,g). This effect persisted after controlling for canonical cortical markers of aging, including hippocampal, superior frontal and association-cortical volumes, and was not observed for these cortical comparators. The same moderation pattern was replicated in UKB executive measures (Fig. 3f), reinforcing that cerebellar integrity supports maintenance of cognitive function across independent samples. Together, these convergent findings align with cerebellar scaffolding frameworks³⁸, in which cerebellar circuits stabilize distributed cortical computations as neocortical systems decline, positioning the cerebellum as a structural substrate of cognitive resilience in later life.

Cerebellar contributions along the AD continuum

Our findings align with accumulating evidence implicating the cerebellum in MCI and AD dementia. Although historically considered relatively spared, the cerebellum exhibits amyloid, tau pathology and microgliosis³⁹, and structural or functional alterations even in the prodromal stages^{14,15,40}. Posterior cerebellar regions aligned with association networks appear particularly vulnerable¹⁵, paralleling network-selective degeneration in the cerebrum; cerebellar structure has been linked to proxies of cognitive reserve in MCI⁴¹.

Stratification according to amyloid status revealed that cerebellar volume was related to cognitive resilience primarily in individuals with a lower amyloid burden (Fig. 4a), which is consistent with a threshold-reserve framework⁴². Cerebellar atrophy was evident across the amyloid spectrum, yet the strongest cognitive associations emerged in patients with clinically diagnosed AD with a lower amyloid burden (Fig. 4c) and in *APOE* ε4/ε4 carriers (Fig. 4d), suggesting that

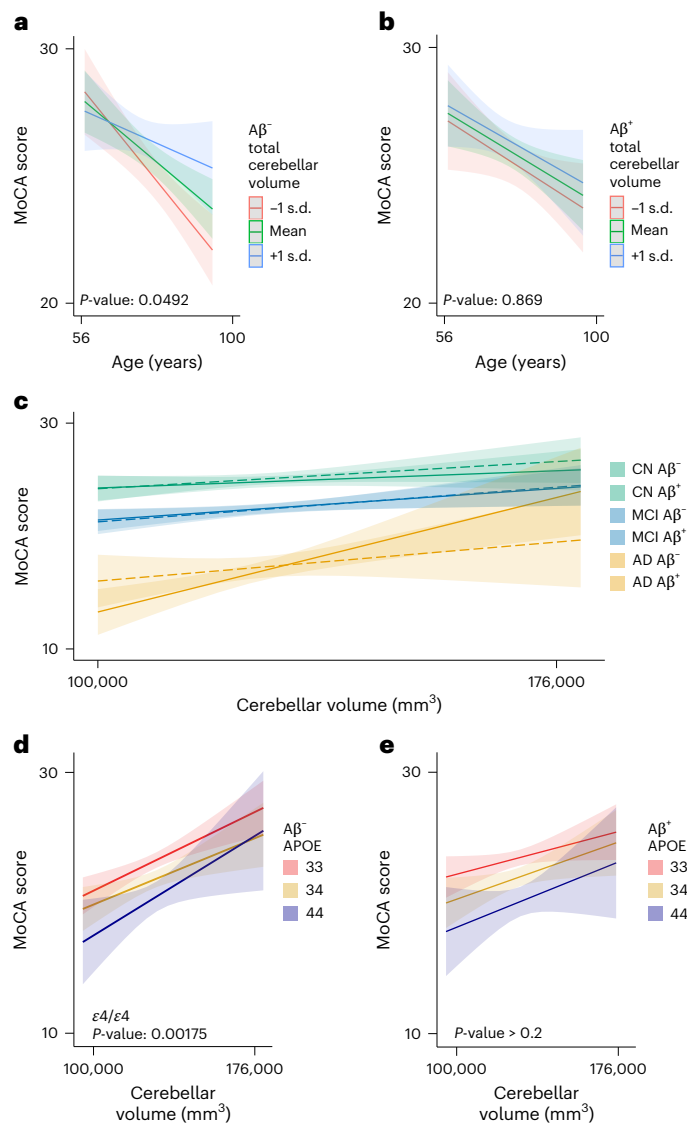


Fig. 4 | Cerebellar volume and MoCA score relationships in ADNI. Analyses were conducted in $n = 1,350$ independent participants from the ADNI cohort with complete data across all variables, with the individual participant as the unit of analysis. Linear models (two-sided tests of interaction terms) controlled for sex and eTIV. **a, b**, Predicted MoCA scores are shown as a function of age at three levels of total cerebellar volume (-1 s.d., mean, $+1$ s.d.), derived from linear interaction models (two-sided) including an age $\times$ cerebellar volume term. The lines represent model-estimated mean predicted MoCA scores; the shaded bands denote the 95% CIs of the regression estimate. **a**, In $A\beta^{-}$ participants, the age $\times$ cerebellar volume interaction was significant ($P = 0.049$). **b**, In $A\beta^{+}$ participants, the interaction was not significant ($P = 0.869$). Axes are scaled to the observed ranges. Lines and bands are model-based predictions. **c**, Stratified linear regression models examined cerebellar volume–MoCA associations according to amyloid and diagnostic status predicting MoCA scores. The lines represent model-estimated mean responses; the shaded bands denote the 95% CIs. In $A\beta^{-}$ individuals (solid line), the association between cerebellar volume and MoCA was strongest among participants with an AD dementia clinical diagnosis (orange) ($P = 0.004$). The color indicates the diagnosis (CN: green, MCI: blue, AD: orange); line type indicates amyloid status (solid: $A\beta^{-}$, dashed: $A\beta^{+}$). **d, e**, Further stratification according to APOE genotype across diagnostic groups ($n = 1,347$ individual participants) revealed that among $A\beta^{-}$ individuals (**d**), cerebellar volume was most strongly associated with MoCA scores in $\epsilon 4/\epsilon 4$ carriers ($P = 0.00175$). No significant $APOE \times$ volume effects were observed in $A\beta^{+}$ individuals (**e**).

cerebellar structural integrity may buffer cognitive symptoms early in the disease trajectory before compensatory capacity is exceeded.

Consistent with this interpretation, cerebellar structure remained relatively stable across diagnostic clinical stages (Supplementary Fig. 7), contrasting with pronounced neocortical atrophy. Moderation effects were most evident in $A\beta^{-}$ individuals, in line with evidence that reserve mechanisms operate preferentially at the earlier stages of pathology⁴³. Together, these findings suggest that cerebellar structural integrity may contribute to cognitive resilience across the continuum from normative aging to early AD.

Genetic and mechanistic considerations

The regional vulnerability observed in this study likely reflects convergent effects of glial physiology, network connectivity and genetic risk. Beyond $APOE \epsilon 4$, AD-associated loci, including *CLU*, *CRI* and *PICALM*, implicate lipid metabolism, endocytosis and microglia-mediated immune pathways⁴⁴. The cerebellum, and posterior lobules, densely interconnected with association cortex, are metabolically demanding and enriched in oligodendrocyte⁴⁵ and microglial populations^{46,47} that may heighten susceptibility. Notably, cerebellar microglia exhibit a transcriptionally distinct, immune-enriched and metabolism-enriched profile and show pronounced age-related shifts relative to other brain regions⁴⁶.

Iron accumulation, a recognized driver of degeneration^{48,49}, may further amplify oxidative stress, microglial activation and myelin disruption. $APOE \epsilon 4$ could exacerbate these processes through impaired lipid homeostasis and heightened neuroinflammatory signaling⁵⁰. Such mechanisms provide a plausible basis for posterior lobular vulnerability and amplified effects in genetic risk carriers.

Longitudinal and methodological context

Although our analyses were cross-sectional, the cerebellar volume–cognition association was replicated across three independent cohorts. Exploratory ADNI analyses further suggested that individuals with larger cerebella maintained better cognitive performance across visits, consistent with a reserve effect. Several limitations warrant consideration. The cross-sectional design limited inference about the temporal relationships between cerebellar atrophy and cognitive decline, and MoCA's global scope limited domain-specific interpretation. Aging effects were spatially complex rather than strictly anterior–posterior, and incomplete education and socioeconomic data limited formal modeling of reserve proxies. Further, future multimodal longitudinal studies integrating structural, functional and biomarker measures will be essential to clarify how cerebellar structure supports cognitive resilience over time.

Implications and conclusions

Across three large cohorts, cerebellar aging showed a reproducible pattern with strongest effects in the posterior lobules and lobule V. Cerebellar volume predicted cognitive resilience in normative aging and AD samples, positioning the cerebellum as an active component of brain aging. Extending reserve frameworks, cerebellar integrity appeared to buffer cognition, particularly before substantial amyloid burden, which is consistent with scaffolding and threshold-reserve models in which structural resources sustain function until pathology overwhelms compensation. Therefore, cerebellar structure may mask emerging symptoms despite advancing cortical pathology, potentially delaying detection under cortex-centric models. Integrating cerebellar metrics into neurodegeneration models, and targeting cerebellar–cortical circuits therapeutically, may strengthen strategies to preserve cognitive health.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-026-02289-x>.

References

- Garmany, A., Yamada, S. & Terzic, A. Longevity leap: mind the healthspan gap. *NPJ Regen. Med.* **6**, 57 (2021).
- King, M., Hernandez-Castillo, C. R., Poldrack, R. A., Ivry, R. B. & Diedrichsen, J. Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nat. Neurosci.* **22**, 1371–1378 (2019).
- Adamaszek, M. et al. Consensus paper: cerebellum and emotion. *Cerebellum* **16**, 552–576 (2017).
- Arleo, A. et al. Consensus paper: cerebellum and ageing. *Cerebellum* **23**, 802–832 (2024).
- Liu, X., d'Oleire Uquillas, F., Viaene, A. N., Zhen, Z. & Gomez, J. A multifaceted gradient in human cerebellum of structural and functional development. *Nat. Neurosci.* **25**, 1129–1133 (2022).
- Raz, N., Dupuis, J. H., Briggs, S. D., McGavran, C. & Acker, J. D. Differential effects of age and sex on the cerebellar hemispheres and the vermis: a prospective MR study. *AJNR Am. J. Neuroradiol.* **19**, 65–71 (1998).
- Andersen, B. B., Gundersen, H. J. G. & Pakkenberg, B. Aging of the human cerebellum: a stereological study. *J. Comp. Neurol.* **466**, 356–365 (2003).
- Raz, N. et al. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cereb. Cortex* **15**, 1676–1689 (2005).
- Han, S., An, Y., Carass, A., Prince, J. L. & Resnick, S. M. Longitudinal analysis of regional cerebellum volumes during normal aging. *Neuroimage* **220**, 117062 (2020).
- Bernard, J. A. & Seidler, R. D. Relationships between regional cerebellar volume and sensorimotor and cognitive function in young and older adults. *Cerebellum* **12**, 721–737 (2013).
- Koppelmans, V. et al. Regional cerebellar volumetric correlates of manual motor and cognitive function. *Brain Struct. Funct.* **222**, 1929–1944 (2017).
- Stern, Y., Barnes, C. A., Grady, C., Jones, R. N. & Raz, N. Brain reserve, cognitive reserve, compensation, and maintenance: operationalization, validity, and mechanisms of cognitive resilience. *Neurobiol. Aging* **83**, 124–129 (2019).
- Schmahmann, J. D. Dysmetria of thought: clinical consequences of cerebellar dysfunction on cognition and affect. *Trends Cogn. Sci.* **2**, 362–371 (1998).
- Jacobs, H. I. L. et al. The cerebellum in Alzheimer's disease: evaluating its role in cognitive decline. *Brain* **141**, 37–47 (2018).
- Guo, C. C. et al. Network-selective vulnerability of the human cerebellum to Alzheimer's disease and frontotemporal dementia. *Brain* **139**, 1527–1538 (2016).
- Mitoma, H. et al. Consensus paper. Cerebellar reserve: from cerebellar physiology to cerebellar disorders. *Cerebellum* **19**, 131–153 (2020).
- King, K. S. et al. Effect of leukocyte telomere length on total and regional brain volumes in a large population-based cohort. *JAMA Neurol.* **71**, 1247–1254 (2014).
- Kerestes, R. et al. Cerebellar volume and disease staging in Parkinson's disease: an ENIGMA-PD study. *Mov. Disord.* **38**, 2269–2281 (2023).
- Glasser, M. F. & Van Essen, D. C. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* **31**, 11597–11616 (2011).
- Preziosa, P. et al. Neurite density explains cortical T1-weighted/T2-weighted ratio in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **92**, 790–792 (2021).
- Burt, J. B. et al. Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nat. Neurosci.* **21**, 1251–1259 (2018).
- Glasser, M. F., Goyal, M. S., Preuss, T. M., Raichle, M. E. & Van Essen, D. C. Trends and properties of human cerebral cortex: correlations with cortical myelin content. *Neuroimage* **93**, 165–175 (2014).
- Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C. & Yeo, B. T. T. The organization of the human cerebellum estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 2322–2345 (2011).
- d'Oleire Uquillas, F. et al. Multimodal evidence for cerebellar influence on cortical development in autism: structural growth amidst functional disruption. *Mol. Psychiatry* **30**, 1558–1572 (2025).
- Zelazo, P. D. The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* **1**, 297–301 (2006).
- Zelazo, P. D. et al. NIH Toolbox Cognition Battery (CB): validation of executive function measures in adults. *J. Int. Neuropsychol. Soc.* **20**, 620–629 (2014).
- Weintraub, S. et al. Cognition Assessment using the NIH Toolbox. *Neurology* **80**, S54–S64 (2013).
- Gershon, R. C. et al. NIH toolbox for assessment of neurological and behavioral function. *Neurology* **80**, S2–S6 (2013).
- Wechsler, D. *Manual for the Wechsler Adult Intelligence Scale* (Psychological Corporation, 1955).
- Reitan, R. M. Validity of the trail making test as an indicator of organic brain damage. *Percept. Mot. Skills* **8**, 271–276 (1958).
- Bernard, J. A. & Seidler, R. D. Moving forward: age effects on the cerebellum underlie cognitive and motor declines. *Neurosci. Biobehav. Rev.* **42**, 193–207 (2014).
- Kim, C.-H. et al. Lobule-specific membrane excitability of cerebellar Purkinje cells. *J. Physiol.* **590**, 273–288 (2012).
- Zhou, H. et al. Cerebellar modules operate at different frequencies. *eLife* **3**, e02536 (2014).
- Sawyer, K. S. et al. Associations between cerebellar subregional morphometry and alcoholism history in men and women. *Alcohol Clin. Exp. Res.* **40**, 1262–1272 (2016).
- Bartzokis, G. et al. In vivo MR evaluation of age-related increases in brain iron. *AJNR Am. J. Neuroradiol.* **15**, 1129–1138 (1994).
- Reisberg, B. et al. in *Neurodevelopment, Aging and Cognition* (eds, Kostovic, I., et al.) 345–369 (Birkhäuser, 1992).
- Reisberg, B., Ferris, S. H. & Franssen, E. H. in *Biological Psychiatry* Vol. 7 (Shagass, C. ed.) 1319–1321 (Elsevier Science Publishing, 1986).
- Bernard, J. A. Don't forget the little brain: a framework for incorporating the cerebellum into the understanding of cognitive aging. *Neurosci. Biobehav. Rev.* **137**, 104639 (2022).
- Sepulveda-Falla, D. et al. Familial Alzheimer's disease-associated presenilin-1 alters cerebellar activity and calcium homeostasis. *J. Clin. Invest.* **124**, 1552–1567 (2014).
- Wang, Z. et al. Spatial patterns of intrinsic brain activity in mild cognitive impairment and Alzheimer's disease: a resting-state functional MRI study. *Hum. Brain Mapp.* **32**, 1720–1740 (2011).
- Devita, M. et al. The role of cognitive reserve in protecting cerebellar volumes of older adults with mild cognitive impairment. *Cerebellum* **23**, 1966–1974 (2024).
- Satz, P. Brain reserve capacity on symptom onset after brain injury: a formulation and review of evidence for threshold theory. *Neuropsychology* **7**, 273–295 (1993).
- Reed, B. R. et al. Measuring cognitive reserve based on the decomposition of episodic memory variance. *Brain* **133**, 2196–2209 (2010).
- Jun, G. et al. Meta-analysis confirms CR1, CLU, and PICALM as Alzheimer disease risk loci and reveals interactions with APOE genotypes. *Arch. Neurol.* **67**, 1473–1484 (2010).

45. Hashimoto, R. et al. Origins of oligodendrocytes in the cerebellum, whose development is controlled by the transcription factor, Sox9. *Mech. Dev.* **140**, 25–40 (2016).
46. Grabert, K. et al. Microglial brain region-dependent diversity and selective regional sensitivities to aging. *Nat. Neurosci.* **19**, 504–516 (2016).
47. Dukhinova, M. S. et al. Cerebellar microglia: on the edge between neuroinflammation and neuroregulation. *Neural Regen. Res.* **21**, 156–172 (2026).
48. Alrouji, M. et al. Iron homeostasis and neurodegeneration in the ageing brain: insight into ferroptosis pathways. *Ageing Res. Rev.* **102**, 102575 (2024).
49. Gao, Q. et al. Role of iron in brain development, aging, and neurodegenerative diseases. *Ann. Med.* **57**, 2472871 (2025).
50. Yin, C. et al. ApoE attenuates unresolvable inflammation by complex formation with activated C1q. *Nat. Med.* **25**, 496–506 (2019).

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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2026, modified publication 2026

Methods

Participants and cohorts

In this cross-sectional study, we sourced MRI data from three cohorts (Table 1) spanning the adult lifespan: the HCP-aging ($n = 708$; aged 36–100), the UKB ($n = 45,013$; aged 44–81) and the ADNI ($n = 1,423$; aged 56–95), totaling 47,144 participants. Sex distributions are reported in Tables 1–3 (HCP-aging, female = 400 (56.5%); UKB, 52.9% female; ADNI sex counts are provided for each diagnostic and amyloid cell in Table 3). All participants provided informed consent and were compensated. Studies were conducted under institutional review board approval.

Demographics

We included participants with complete cerebellar MRI coverage and cognitive testing. In HCP-aging, there were two (0.28%) American Indian or Alaskan Native individuals, 52 (7.20%) Asian, 98 (13.84%) Black or African American, 506 (71.61%) White, 36 (4.38%) with more than one ethnicity and 14 (1.98%) individuals with unknown or unreported ethnicity, with 69 (9.75%) identifying as Hispanic or Latino, 632 (89.27%) not Hispanic nor Latino and seven (0.28%) of unreported ethnic group. We also sourced MRI data from the UKB but did not have access to ethnicity or ethnic group status. Additionally, the ADNI cohort consisted of two (0.14%) American Indian or Alaskan Native individuals, seven (0.50%) Asian, 22 (1.55%) Black or African American, one (0.07%) Native Hawaiian or other Pacific Islander, 670 (47.08%) White, 11 (0.77%) individuals of more than one ethnicity and 710 (49.89%) individuals of unknown or unreported ethnic status, with 16 (1.12%) identifying as Hispanic or Latino, 694 (48.77%) not Hispanic or Latino and 713 (50.11%) of unreported ethnic group.

Clinical characterization (ADNI)

For ADNI participants, a clinical diagnosis of MCI or AD was ascertained by a consensus of trained clinicians, considering cognitive scores including MoCA⁵¹, the Clinical Dementia Rating Scale⁵² (score of 0 for CN function, 0.5 for MCI, 1 for mild dementia, 2 for moderate dementia and 3 for severe and advanced dementia) and the Alzheimer's Disease Inventory. Our ADNI analysis considered a total of 565 Clinical Dementia Rating (CDR) = 0, 672 CDR = 0.5, 147 CDR = 1, 33 CDR = 2 and four CDR = 3 individuals. The ADNI sample was stratified according to a CN, MCI and AD dementia clinical diagnosis. We opted to include individuals with a 'subjective cognitive concerns' in the CN group, given previous reports showing that optimal subgroup classification of ADNI data using spectral embedding, multidimensional scaling, uniform manifold approximation and projection and t -distributed stochastic neighbor embedding, is best captured by the CN, MCI and AD diagnostic groupings⁵³.

Cognitive assessments

Cognitive assessments were tailored based on the cohort; they included the MoCA⁵¹ within 365 days of the MRI procedures, the DSST²⁹ and TMB³⁰, tests of processing speed and attention, and the Clinical Dementia Rating Scale⁵². No MoCA cutoff was imposed for inclusion. Among the HCP-aging participants included in the analyses, MoCA ranged from 19 to 30 (Q1 = 25, median = 27, mean = 26.32, Q3 = 28). This wide variability is consistent with a community sample without a MoCA-based inclusion threshold.

MoCA scores below 26 in the HCP-aging cohort reflect expected variability in community samples, as optimal screening cutoffs differ across cohorts (for example, Nasreddine et al.⁵¹). Nonetheless, lower scores may include individuals with an early or prodromal pathology. Similarly, A β -yet cognitively impaired ADNI participants may represent non-AD conditions (for example, suspected non-AD pathology), with heterogeneous trajectories⁵⁴. Although amyloid-stratified analyses mitigate this concern, we interpret cerebellar–cognition associations conservatively and do not infer causality.

MRI and positron emission tomography acquisition

All participants underwent 3T T1-weighted MRI. Structural images were segmented using FreeSurfer (v.7.1.1) and ACAPULCO⁵⁵ for cerebellar parcellation in HCP and ADNI, and FMRIB software library-derived morphometry for the UKB (see Miller et al.⁵⁶ and Alfaro-Almagro et al.⁵⁷ for processing details). The cortex was parcellated using the HCP-MMP atlas¹⁹.

In ADNI, amyloid-PET [⁶¹C]-Pittsburgh compound B was acquired 50–70 min after injection and processed centrally to derive partial-volume-corrected regional uptake value ratios. The *APOE* genotype was determined from blood samples using standard assays. Additional acquisition parameters are provided in Supplementary Table 9.

Statistical analysis

We quantified associations between sample characteristics using R (v.4.0.4). Multivariate linear regression models were conducted across the HCP, UKB and ADNI cohorts, covarying for biological sex, eTIV (from FreeSurfer) or total brain volume (from the FMRIB software library), and years of education for the HCP and the UKB samples. Robust standard error estimates were calculated unless otherwise specified; standardized β -coefficients are reported for all models along with Cohen's f effect sizes for partial effects in multivariate regression.

Lobule-wise aging models (HCP)

In the HCP cohort, we investigated relationships between age, total cerebellar volume (ACAPULCO parcellation), total neocortical volume and eTIV from FreeSurfer, and MoCA scores. Lobular gray matter volumes (mm³) were log-transformed and modeled using mixed-effects models with age $\times$ lobule interactions, adjusting for sex, eTIV and years of education.

Primary analyses used bilateral summed homologous lobules, defined a priori to (1) increase signal to noise, (2) reduce multiple comparisons and (3) align with our focus on lobule-level aging gradients rather than hemispheric lateralization.

Type III ANOVA tables were computed using Satterthwaite approximations for the numerator and denominator degrees of freedom and are reported as F (numerator d.f., denominator d.f.). All tests were two-sided. Lobule-specific age slopes were estimated on the log scale and converted to percentage change per decade as $100 \times (\exp(10\beta) - 1)$. CIs used asymptotic (z) inference. Left–right slope differences were tested on the log scale and corrected using Benjamini–Hochberg FDR ($q < 0.05$). Hemisphere-specific age $\times$ lobule $\times$ hemisphere models (Supplementary Fig. 1) were performed analogously.

Vermis analyses (HCP)

Supplementary analyses examined the mid-sagittal vermis regions (ACAPULCO: Vermis VI–X). Log-transformed volumes were modeled using linear mixed-effects models with age $\times$ vermis ROI interactions, adjusting for sex, eTIV and years of education.

A Type III ANOVA with Satterthwaite degrees of freedom tested the age $\times$ vermis ROI interaction. Vermis-specific age slopes were estimated and converted to percentage change per decade as $100 \times (\exp(10\beta) - 1)$. We report point estimates and 95% asymptotic (z -based) CIs and provide Benjamini–Hochberg-adjusted pairwise slope comparisons. Percentage-per-decade estimates are visualized in Supplementary Fig. 2 and reported in Supplementary Table 2.

Microstructure analyses: T1W/T2W (HCP)

To index cerebellar tissue microstructure, we analyzed voxel-wise T1W/T2W ratio maps, a unitless contrast sensitive to myelin, neurite content and partial-volume effects content^{19,20}. In the cerebellum, this measure is influenced by tissue microstructural content and intra-voxel cerebrospinal fluid admixture. Ratios were computed within ACAPULCO-defined lobules. We hypothesized reduced age effects in sensorimotor lobules and stronger decline in posterior association lobules.

Log-transformed T1W/T2W values were modeled using linear mixed-effects models with age $\times$ lobule interactions, adjusting for sex, eTIV, years of education and hemisphere, with subject random intercepts. Type III tests with Satterthwaite degrees of freedom provided omnibus inference. Lobule-specific age slopes were estimated and expressed as percentage change per decade: $100 \times (\exp(10\beta) - 1)$. Pairwise slope differences were Benjamini–Hochberg FDR-corrected.

To compare cerebellar and cortical tissue aging patterns, we analyzed 320 cortical ROIs from the HCP Multimodal Parcellation (HCP-MMP1¹⁹) using HCP-aging T1W/T2W maps. For each ROI, age-related slopes were estimated with linear models, and mean T1W/T2W from the HCP young-adult reference group (aged 22–35 years) indexed baseline tissue properties. A general linear model was used to test whether baseline T1W/T2W predicted age-related decline across ROIs. The fitted regression and 95% CIs are shown in Fig. 1i.

qMRI validation cohort

To validate the T1W/T2W findings, qMRI data were acquired from 23 healthy adults (aged 29–71) at the Princeton Neuroscience Institute (3T Siemens Prisma scanner, 64-channel head coil). Multi-flip angle fast low-angle shot three-dimensional scans (FLASH) were acquired with a repetition time of 19 ms, echo time of 3.34 ms and flip angles 4, 10 and 20. Echo-planar inversion recovery scans had a repetition time of 2,920 ms and inversion times of 2,400, 1,200, 400 and 200 ms. For each participant, FLASH and EPI images were processed using a MATLAB pipeline (mrQ⁵⁸), including gradient nonlinearity correction, bias field correction and co-registration to individual T1 images. We analyzed the VIP, a quantitative index of water protons interfacing with cellular structures.

Proton-density-derived VIP maps were normalized to Montreal Neurological Institute space, parcellated with ACAPULCO, and averaged within lobules (normalization visually verified). Age was mean-centered (years). Sex was included as a covariate. Because VIP reflects voxel-wise tissue properties rather than regional volume, intracranial volume was not modeled. Log-transformed VIP values were analyzed using linear mixed-effects models with age $\times$ lobule interactions and subject random intercepts. Type III *F*-tests with Satterthwaite degrees of freedom provided omnibus inference. Lobule-specific age slopes were estimated and expressed as percentage change per decade: $100 \times (\exp(10\beta) - 1)$, with Benjamini–Hochberg FDR correction across lobules. A planned contrast compared posterior cognitive lobules (crus I and II) against motor lobules (I–III, IV), using balanced weights on age slopes (averaged over the hemisphere). Model diagnostics showed approximately normal residuals and no singular fits.

qMRI measures were derived to characterize cerebellar microstructure beyond volume. These included TV, indexing macromolecular tissue content, WTF (WTF = 1 – TV) and VIP, a composite metric of tissue microstructure after accounting for free water. Because TV and WTF explicitly model tissue and water contributions, they are less susceptible to cerebrospinal fluid-related partial-volume effects, particularly in thin folia. VIP builds on this framework and is expected to be more robust than heuristic ratios such as T1W/T2W. For supplementary analyses, lobular qMRI measures were averaged across hemispheres, and age associations were tested using linear models correcting for sex.

Lobule-wise cognition models (HCP)

To assess the relationship between cognition and cerebellar volume at the lobular level, bilaterally summed ACAPULCO volumes were log-transformed and entered into multiple regression models predicting MoCA. Covariates included sex, eTIV and years of education. Coefficients were rescaled to reflect Δ MoCA per 10% increase in volume. 95% CIs were derived from the model standard errors.

Pairwise slope differences were tested using Wald contrasts from the model's variance–covariance matrix and corrected using Benjamini–Hochberg FDR. Guided by cerebellar functional organization,

we contrasted posterior association lobules (crus I and crus II) with anterior and motor lobules (I–III, IV, V) using equal-weight linear combinations, reported as Δ MoCA per 10% volume with CIs.

VBM analyses (HCP)

VBM was performed in CAT12⁵⁹ (<https://neuro-jena.github.io/cat/>). CAT12 uses a standardized pipeline optimized for VBM analyses in SPM12⁶⁰, including bias correction of field intensity non-uniformities and tissue segmentation with partial-volume modeling. Analyses used modulated, normalized maps (nonlinear modulation only), with eTIV as a covariate. Statistical conjunction of corrected maps was performed in FreeSurfer to identify regions associated with age and cognition (FWE-corrected $P < 0.05$).

In supplementary analyses, voxel-wise associations between cerebellar gray matter and individual MoCA subscales were tested in CAT12/SPM, adjusting for eTIV (Supplementary Fig. 3). Maps are displayed as (positive) signed values on the SUIT flatmap⁶¹, with green contours marking $P < 0.001$ (uncorrected) (Supplementary Fig. 3). The group-level voxel-wise SPM *T*-map shows reduced coverage in inferior cerebellar regions. This reflects conservative group-level masking during CAT12/SPM second-level estimation (for example, implicit intersection masking across individuals and tissue probability thresholds), rather than insufficient cerebellar coverage in the underlying data. Visual inspection of individual-level scans and cerebellar parcellations confirmed appropriate inclusion and segmentation of inferior cerebellar lobules across participants.

Cortical comparators and reserve modeling (HCP)

Cortical comparators included bilateral Desikan–Killiany composites for the PPCC/DMN (precuneus and posterior cingulate), FPN (rostral middle frontal, caudal middle frontal, inferior parietal) and VAN (insula, pars opercularis, pars triangularis, supramarginal). Gray matter volumes (mm³) were z-scored within sample. Age was mean-centered. Covariates included sex, eTIV and years of education.

Reserve was tested via age $\times$ volume interactions predicting MoCA. Single-moderator models included one ROI at a time. Dual-moderator models included both the cerebellar MoCA signature and a cortical ROI with corresponding interaction terms. We report interaction coefficients, *t* statistics and 95% CIs. Moderation effects were compared directly using Wald tests, with Benjamini–Hochberg correction across related tests.

To test combinatorial reserve effects, we fitted linear models including age $\times$ cerebellum $\times$ cortex (cortex = PPCC/DMN, FPN or VAN composites), with all lower-order terms and covariates (sex, eTIV, years of education). The three-way interaction term indexed age-slope modulation according to regional volume. Robustness was evaluated using HC3 heteroskedasticity-consistent standard errors and Benjamini–Hochberg correction across cortical ROIs.

Specificity analyses examined bilateral hippocampal and superior frontal volumes (Desikan–Killiany). We tested age–volume associations, volume–MoCA associations, age–volume interactions on MoCA and whether cerebellar volume $\times$ age effects persisted after controlling for hippocampal or superior frontal volumes (Supplementary Fig. 4). All models included sex, eTIV and years of education as covariates. Interaction effects were visualized using simple slopes at -1 s.d., mean and $+1$ s.d. of the moderator.

Sex interaction analyses (HCP)

We also tested sex interactions with key predictors using prespecified linear models. MoCA – age $\times$ sex + eTIV + years of education; MoCA – years of education $\times$ sex + eTIV; MoCA – neocortical volume $\times$ sex + eTIV + years of education; MoCA – eTIV $\times$ sex + years of education; MoCA – cerebellar volume $\times$ sex + eTIV + years of education (Supplementary Fig. 5). For the voxel-wise cerebellar VBM gray matter analyses, we repeated age-related and MoCA-related models stratified according to sex (with eTIV as

covariate); the age effect VBM map survived correction for multiple comparisons (Supplementary Fig. 5g, top) (FWE-corrected $P < 0.05$), while the MoCA effect VBM map did not and is shown in Supplementary Fig. 5g (bottom) as uncorrected at $P < 0.0001$.

Task-level validation (HCP)

In HCP-aging, we tested whether the aging-MoCA cerebellar ROI moderates age-related differences in NIH Toolbox performance: dimensional card sorting task²⁵, flanker inhibitory control & attention task²⁶, PSM²⁹ and LSWM^{28,62}. For each task, we fitted ordinary least-squares linear models: score = age $\times$ cerebellar volume + sex + eTIV + years of education, with primary inference on the age $\times$ cerebellar volume interaction (Supplementary Fig. 6). Interaction effect sizes are reported as Cohen's f . Significant interactions were probed using simple-slope/Johnson–Neyman procedures (two-sided and FDR-adjusted within the task). To assess additive structure–performance associations, we also fitted models without the interaction term.

UKB validation analyses

In the UKB validation sample, we examined associations between cerebellar volume and age, and cognitive performance (DSST²⁹; TMB³⁰ tests). Lobule-specific aging effects were tested for crus I, crus II and lobule IX (left/right hemispheres retained) using mixed-effects models. Lobular volumes were log-transformed to estimate percentage change per decade: $100 \times (\exp(10\beta^*) - 1)$.

Type III tests with Satterthwaite degrees of freedom provided omnibus inference. Lobule-specific age slopes were estimated and expressed as percentage per decade with 95% CIs; pairwise slope differences were examined with Benjamini–Hochberg correction for multiple comparisons. Gaussian mixture models²⁰ classified participants into high-volume and low-volume subgroups based on mean crus I and crus II volume.

Age, cerebellar volume and cognition across amyloid and APOE (ADNI)

In ADNI, we examined associations among cerebellar and neocortical volumes, age, APOE genotype and MoCA scores in CN participants, and participants with MCI and AD dementia. Analyses were first stratified according to amyloid status ($A\beta^-/A\beta^+$), testing the moderation of the cross-sectional age–MoCA association according to total cerebellar volume: MoCA = age $\times$ cerebellar volume + sex + eTIV. Education or socioeconomic status were not included because of high missingness or unavailability. Predicted lines were generated at -1 s.d., mean and $+1$ s.d. of cerebellar volume (95% CIs; two-sided).

We then fitted three hypothesis-driven models using diagnostic group according to amyloid status (CN $A\beta^-$, CN $A\beta^+$, MCI $A\beta^-$, MCI $A\beta^+$, AD $A\beta^-$, AD $A\beta^+$): (1) MoCA = cerebellar volume $\times$ group; (2) MoCA = age $\times$ cerebellar volume $\times$ group; (3) MoCA = cerebellar volume $\times$ APOE zygosity. All adjusting for age, sex and eTIV. Given the limited number of a priori models, P values are reported unadjusted. Sample sizes were: 303 CN $A\beta^-$; 280 CN $A\beta^+$; 355 MCI $A\beta^-$; 336 MCI $A\beta^+$; 80 AD $A\beta^-$; and 62 AD $A\beta^+$. In the $A\beta^-$ group, there were 352 APOE 3/3 carriers, 252 3/4 carriers and 54 4/4 carriers. In the $A\beta^+$ group, there were 331 3/3 carriers, 214 3/4 carriers and 65 4/4 carriers.

Lobule-wise age and cognition effects (ADNI)

Within the $A\beta^-$ and $A\beta^+$ strata, we fitted linear mixed-effects models with individual random intercepts to test hemispheric lobule-specific aging effects: $\log(\text{volume}) = \text{age} \times \text{lobule} \times \text{hemisphere} + \text{sex} + \text{eTIV}$. Age slopes were estimated and expressed as percentage change per decade $100 \times (\exp(10\beta^*) - 1)$, with 95% CIs and Benjamini–Hochberg-adjusted pairwise comparisons.

To examine lobule-specific volume–cognition associations, we modeled MoCA as a function of log-transformed lobule volume (separately within $A\beta^-$ and $A\beta^+$), adjusting for covariates (Supplementary Fig. 8).

Effects were expressed as ΔMoCA per 10% greater lobule volume, with 95% CIs and Benjamini–Hochberg correction.

Longitudinal cognition models (ADNI)

To test whether baseline cerebellar volume predicts longitudinal cognitive change, we fitted linear mixed-effects models using all available MoCA assessments across visits for every individual. Models included time since baseline (years; mean-centered per participant), baseline cerebellar volume, age, sex and eTIV, with participant as a random intercept. Random slope models additionally included participant-specific slopes for time: MoCA = time $\times$ cerebellar volume + age + sex + eTIV + (1|individual) + (0 + time|individual). We also evaluated a simpler random-intercept-only model. Baseline MoCA was not included as a covariate given evidence that repeated cognitive testing can lead to practice-related score improvements that do not reflect true cognitive change^{63,64}. Therefore, including baseline MoCA can obscure the genuine longitudinal relationship between brain structure and cognitive trajectory by absorbing interindividual variance that partly reflects learning effects. Finally, we note that individuals who remain in longitudinal studies longer often represent a healthier survivor subgroup^{65,66}, which can attenuate brain–behavior associations over time.

Reporting summary

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

Data availability

Data analyzed are available at public repositories, which may require appropriate data sharing agreements from each individual publicly accessible database, including LONI for ADNI and HCP data (<https://ida.loni.usc.edu/>), and biobank UK for UKB data (<https://biobank.ctsu.ox.ac.uk>). Deidentified data for Figs. 1–4 can be found at <https://github.com/fdoleireuquillas/AgingCerebellum>. Source data are provided with this paper.

Code availability

Source code related to the statistical analysis and visualization can be found at <https://github.com/fdoleireuquillas/AgingCerebellum>.

References

- Nasreddine, Z. S. et al. in *Montreal Cognitive Assessment (MoCA)* (APA, 2005).
- Morris, J. C. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* **43**, 2412–2414 (1993).
- van der Haar, D., Moustafa, A., Warren, S. L., Alashwal, H. & van Zyl, T. An Alzheimer's disease category progression sub-grouping analysis using manifold learning on ADNI. *Sci. Rep.* **13**, 10483 (2023).
- Wisse, L. E. M. et al. Defining SNAP by cross-sectional and longitudinal definitions of neurodegeneration. *Neuroimage Clin.* **18**, 407–412 (2018).
- Han, S., Carass, A., He, Y. & Prince, J. L. Automatic cerebellum anatomical parcellation using U-Net with locally constrained optimization. *Neuroimage* **218**, 116819 (2020).
- Miller, K. L. et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).
- Alfaro-Almagro, F. et al. Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* **166**, 400–424 (2018).
- Mezer, A. et al. Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nat. Med.* **19**, 1667–1672 (2013).
- Gaser, C. et al. CAT: a computational anatomy toolbox for the analysis of structural MRI data. *Gigascience* **13**, giae049 (2024).

60. Ashburner, J. & Friston, K. J. Unified segmentation. *Neuroimage* **26**, 839–851 (2005).
61. Diedrichsen, J. & Zotow, E. Surface-based display of volume-averaged cerebellar imaging data. *PLoS ONE* **10**, e0133402 (2015).
62. McLachlan, G. J., Lee, S. X. & Rathnayake, S. I. Finite mixture models. *Annu. Rev. Stat. Appl.* **6**, 355–378 (2019).
63. Goldberg, T. E., Harvey, P. D., Wesnes, K. A., Snyder, P. J. & Schneider, L. S. Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. *Alzheimers Dement.* **1**, 103–111 (2015).
64. Calamia, M., Markon, K. & Tranel, D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *Clin. Neuropsychol.* **26**, 543–570 (2012).
65. Nyberg, L., Lövdén, M., Riklund, K., Lindenberg, U. & Bäckman, L. Memory aging and brain maintenance. *Trends Cogn. Sci.* **16**, 292–305 (2012).
66. Lindenberg, U. Human cognitive aging: corriger la fortune? *Science* **346**, 572–578 (2014).

Acknowledgements

We thank V. Zhang and J. Merriman for their assistance with the quality control of the HCP dataset during the early phases of this project. We also thank N. Daw for valuable discussions and conceptual input during early meetings that helped shape the direction of this work. This study was supported by funding from the National Science Foundation through a Graduate Research Fellowship (F.d.U.: no. 2021314805) and a CAREER Award (J.G.: no. 2337373), and the National Academies of Sciences, Engineering, and Medicine through a Ford Foundation Predoctoral Fellowship Award (F.d.U.: no. 20210323). Funding for the acquisition of the data was provided in part by the Regina and John Scully '66 Center for the Neuroscience of Mind and Behavior. We also acknowledge funding to J.D.C. from the Department of Defense Office of Naval Research Vannevar Bush Faculty Fellowship (N00014-22-12002). Data used in preparation of this article were obtained from the ADNI database (<https://adni.loni.usc.edu/>). As such, investigators within ADNI contributed to the design and implementation of ADNI or provided data but did not participate in the analysis or the writing of this report. The funders had no role in study design, data collection, analysis, interpretation or manuscript preparation. Data were partially obtained from ADNI, whose investigators contributed to data acquisition but not to data analysis or writing. ADNI was launched in 2003 and led by M. W. Weiner to test whether MRI, PET, biomarkers and clinical assessments can track progression of mild cognitive impairment and AD. Data collection and sharing for this project was funded by ADNI (National Institutes of Health Grant no. U01 AG024904) and DOD ADNI (Department of Defense award no. W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, The National Institute of Biomedical

Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica; Biogen; Bristol Myers Squibb; CereSpir; Cogstate; Eisai; Elan Pharmaceuticals; Eli Lilly and Company; EuroImmun; F. Hoffmann–La Roche Ltd and its affiliated company Genentech; Fujirebio; GE Healthcare; IXICO; Janssen Alzheimer Immunotherapy Research & Development.; Johnson & Johnson Pharmaceutical Research & Development; Lumosity; Lundbeck; Merck & Co.; Meso Scale Diagnostics; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (<https://fnih.org/>). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Author contributions

F.d.U. and J.G. conceptualized and designed the study. F.d.U., A.F.A.-B., E.D.H., J. Seidlitz, P.V., R.A.I.B., R.R.-G. and V.W. contributed to data access, curation and processing. F.d.U., E.D.H., E.S., J.G., J. Seidlitz, P.V. and S.S.-H.W. contributed to data analysis. F.d.U., E.D.H., E.S., J.D.C., J.G., J. Seidlitz, J. Sepulcre, P.V. and S.S.-H.W. contributed to data interpretation and guided study progress. F.d.U., J.G., P.V. and S.S.-H.W. were responsible for writing the manuscript.

Competing interests

J. Seidlitz and R.A.I.B. are directors of and hold equity in Centile Biosciences. A.F.A.-B. sits on the advisory board and holds equity in Centile Biosciences. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41593-026-02289-x>.

Correspondence and requests for materials should be addressed to Federico d'Oleire Uquillas.

Peer review information *Nature Neuroscience* thanks Jessica Bernard, Jorn Diedrichsen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| ☐ | ☒ | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Siemens 3T Magnetom Prisma for quantitative MRI data at the Princeton Neuroscience Institute (<https://github.com/mezera/mrQ>)

Data analysis FreeSurfer (7.1.1), ACAPULCO cerebellar segmentation pipeline (Han et al., 2020), R (v.4.0.4: <https://www.R-project.org/>), FSL (6.0.2)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data analyzed are available at public repositories, which may require appropriate data sharing agreements from each individual publicly accessible database, including LONI for ADNI and HCP data (<https://ida.loni.usc.edu/>), and biobankUK for UK Biobank data (<https://biobank.ctsu.ox.ac.uk>). Deidentified source data for main figures can be found at:

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

In HCP-Aging, Female = 400 (56.5%) (Table 1a); in UK Biobank, 52.9% were female (Table 1b); and ADNI sex counts are provided for each diagnostic/amyloid cell (Table 1c). Clinically Normal Amyloid Negative, Female = 55.4%; Clinically Normal Amyloid Positive, Female = 56.1%; Mild Cognitive Impairment Amyloid Negative, Female = 40%; Mild Cognitive Impairment Amyloid Positive, Female = 36.6%; Alzheimer's Disease Dementia Amyloid Negative, Female = 37.5; Alzheimer's Disease Dementia Amyloid Positive, Female = 45.2%. All participants provided informed consent, and each cohort adhered to institutional review board guidelines.

Reporting on race, ethnicity, or other socially relevant groupings

Demographics
We included a racially diverse sample of participants with complete cerebellar coverage on MRI scans, as well as corresponding cognitive testing. In the HCP-Aging cohort, there were 2 (0.28%) American Indian/Alaskan Native individuals, 52 (7.20%) Asian, 98 (13.84%) Black or African American, 506 (71.61%) White, 36 (4.38%) more than one race, and 14 (1.98%) individuals with unknown or unreported race, with 69 (9.75%) identifying as Hispanic or Latino, 632 (89.27%) not Hispanic or Latino, and 7 (0.28%) of unreported ethnic group. We also sourced MRI data from the UK Biobank, but did not have access to ethnic or racial status. Additionally, our ADNI cohort comprised of 2 (0.14%) American Indian or Alaskan Native individuals, 7 (0.50%) Asian, 22 (1.55%) Black or African American, 1 (0.07%) Native Hawaiian or other Pacific Islander, 670 (47.08%) White, 11 (0.77%) individuals of more than one race, and 710 (49.89%) individuals of unknown or unreported racial status, with 16 (1.12%) identifying as Hispanic or Latino, 694 (48.77%) not Hispanic or Latino, and 713 (50.11%) of unreported ethnic group.

Population characteristics

The HCP-Aging cohort included 708 participants aged 36–100 years. The UK Biobank dataset provided 45,013 participants aged 44–81 years, and the ADNI cohort included 1,423 participants aged 56–95 years old.

Recruitment

Human Connectome Project-Aging (HCP-Aging)
Participants were recruited as part of the Human Connectome Project in Aging, a multisite initiative supported by the NIH. Recruitment targeted cognitively healthy adults aged 36–100 years. Individuals were identified through community outreach, registries, family members of HCP participants, and institutional participant pools. Eligibility was determined through structured screening that excluded major neurological or psychiatric illness, contraindications to MRI, and significant cognitive impairment. All participants provided written informed consent under protocols approved by the Washington University in St. Louis Institutional Review Board.

UK Biobank
Participants were recruited through the UK National Health Service population registry. ~500,000 adults aged 40–69 years were invited between 2006–2010 via mailed invitations; those who volunteered attended one of 22 assessment centers throughout the United Kingdom. The neuroimaging subset was recruited from this broader cohort and invited back for MRI scanning based on geographic proximity to imaging centers and willingness to participate in follow-up assessments. Inclusion/exclusion criteria were minimal, consistent with UK Biobank's population-level design, but excluded individuals unable to provide informed consent. Ethical approval was obtained from the UK Biobank Research Ethics Committee, and all participants gave written informed consent.

Alzheimer's Disease Neuromaging Initiative (ADNI)
Participants were recruited across ~60 North American sites using a combination of clinician referral, community outreach, memory clinics, and media advertisements. ADNI enrolls individuals spanning normal aging, mild cognitive impairment (MCI), and Alzheimer's disease (AD) dementia, following prespecified diagnostic criteria based on clinical evaluation, neuropsychological testing, and functional assessment. Key exclusion criteria included significant neurological disease other than AD, major psychiatric disorders, MRI contraindications, and unstable medical illness. Participants provided written informed consent, and study procedures were approved by local institutional review boards at all participating centers.

Princeton validation sample
The Princeton qMRI sample was recruited locally through campus and community advertisements. Eligible participants were healthy adults with no history of neurological or major psychiatric illness and no MRI contraindications. All participants provided written informed consent under a protocol approved by the Princeton University Institutional Review Board.

Ethics oversight

All datasets received IRB/ethics approval and written informed consent (HCP-Aging, UK Biobank, ADNI, Princeton qMRI). Only de-identified data were used.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were determined by the availability of eligible participants within each cohort (HCP-Aging n=708, UK Biobank n=47,144, ADNI n=1,423, Princeton Neuroscience Institute n=23). These are large, well-established neuroimaging studies with prespecified recruitment targets and publicly defined protocols. No a priori power calculations were performed by the present study, as analyses used all participants meeting inclusion/exclusion criteria, imaging quality thresholds, and data completeness requirements.
Data exclusions	Participants were excluded only if imaging data did not meet quality-control criteria. MRI exclusions included scans with insufficient field of view (e.g., incomplete coverage of the cerebellum), substantial motion or artifact, or failed preprocessing. For analyses using ACAPULCO cerebellar parcellation, participants were additionally excluded if automated lobular segmentations did not meet visual quality-control standards. Only individuals with complete demographic data and usable MRI after QC were included.
Replication	Key findings were replicated across independent datasets and imaging modalities. Primary analyses conducted in the HCP-Aging cohort were conceptually and quantitatively replicated in the UK Biobank using analogous structural MRI measures and identical statistical models. We further replicated the spatial pattern of age-related cerebellar differences in an independent quantitative MRI dataset collected at Princeton University, demonstrating consistency across acquisition platforms and imaging modalities. All replication analyses were conducted independently and reproduced the direction and spatial distribution of the original effects.
Randomization	Participants were allocated to groups based on cohort-defined diagnostic or demographic categories rather than random assignment. Since the study was observational, randomization was not applicable. Statistical models included Sex, total cranial/brain volume, and education (when available) to control for systematic group differences.
Blinding	Blinding was not applicable because this was an observational neuroimaging study with no experimentally assigned groups. Imaging preprocessing and statistical analyses were performed using standardized, automatic pipelines. For steps requiring manual quality control (e.g., verifying cerebellar field-of-view coverage or ACAPULCO cerebellar segmentation), raters were blinded to participants' demographic and cognitive data.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<i>Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).</i>
Research sample	<i>State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.</i>
Sampling strategy	<i>Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.</i>
Data collection	<i>Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.</i>
Timing	<i>Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Non-participation	<i>State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.</i>
Randomization	<i>If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.</i>

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<i>Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.</i>
Research sample	<i>Describe the research sample (e.g. a group of tagged <i>Passer domesticus</i>, all <i>Stenocereus thurberi</i> within Organ Pipe Cactus National</i>

Research sample	Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.
Sampling strategy	Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.
Data collection	Describe the data collection procedure, including who recorded the data and how.
Timing and spatial scale	Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken
Data exclusions	If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.
Reproducibility	Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.
Randomization	Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.
Blinding	Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).
Location	State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).
Access & import/export	Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).
Disturbance	Describe any disturbance caused by the study and how it was minimized.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.
Validation	Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.
Authentication	Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.
Mycoplasma contamination	Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology and Archaeology

Specimen provenance	Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.
Specimen deposition	Indicate where the specimens have been deposited to permit free access by other researchers.
Dating methods	If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.
☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.	
Ethics oversight	Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.
Wild animals	Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.
Reporting on sex	Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.
Field-collected samples	For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.
Ethics oversight	Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.
Study protocol	Note where the full trial protocol can be accessed OR if not available, explain why.
Data collection	Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

- | No | Yes | |
|--------------------------|--------------------------|----------------------------|
| ☐ | ☐ | Public health |
| ☐ | ☐ | National security |
| ☐ | ☐ | Crops and/or livestock |
| ☐ | ☐ | Ecosystems |
| ☐ | ☐ | Any other significant area |

Experiments of concern

Does the work involve any of these experiments of concern:

- | No | Yes | |
|--------------------------|--------------------------|---|
| ☐ | ☐ | Demonstrate how to render a vaccine ineffective |
| ☐ | ☐ | Confer resistance to therapeutically useful antibiotics or antiviral agents |
| ☐ | ☐ | Enhance the virulence of a pathogen or render a nonpathogen virulent |
| ☐ | ☐ | Increase transmissibility of a pathogen |
| ☐ | ☐ | Alter the host range of a pathogen |
| ☐ | ☐ | Enable evasion of diagnostic/detection modalities |
| ☐ | ☐ | Enable the weaponization of a biological agent or toxin |
| ☐ | ☐ | Any other potentially harmful combination of experiments and agents |

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.

ChIP-seq

Data deposition

- ☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session (e.g. [UCSC](#))

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates	<i>Describe the experimental replicates, specifying number, type and replicate agreement.</i>
Sequencing depth	<i>Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.</i>
Antibodies	<i>Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.</i>
Peak calling parameters	<i>Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.</i>
Data quality	<i>Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.</i>
Software	<i>Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.</i>

Flow Cytometry

Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	<i>Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.</i>
Instrument	<i>Identify the instrument used for data collection, specifying make and model number.</i>
Software	<i>Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.</i>
Cell population abundance	<i>Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.</i>
Gating strategy	<i>Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.</i>
☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.	

Magnetic resonance imaging

Experimental design

Design type	Structural MRI (T1-weighted scans)
Design specifications	NA
Behavioral performance measures	NA

Acquisition

Imaging type(s)	https://www.humanconnectome.org/hcp-protocols-ya-3t-imaging , https://github.com/mezera/mrQ , https://community.ukbiobank.ac.uk/hc/en-gb/articles/24618819821981-Imaging-Data
Field strength	3T
Sequence & imaging parameters	See links above for parameters
Area of acquisition	Head

Diffusion MRI

☐ Used☒ Not used

Preprocessing

Preprocessing software	See links above in acquisition for pipeline information
Normalization	If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.
Normalization template	Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.
Noise and artifact removal	Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).
Volume censoring	Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings	Offline models of volumetry and age, two sided.
Effect(s) tested	Effect of Age on neocortical volume, cerebellar volume, and cognition. Moderation effect of cerebellar volume on age-cognition associations.
Specify type of analysis:	☐ Whole brain ☐ ROI-based ☒ Both
Statistic type for inference	Both voxel-wise and cluster correction in VBM maps.
(See Eklund et al. 2016)	
Correction	FWE, FDR

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis
Functional and/or effective connectivity	Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).
Graph analysis	Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).
Multivariate modeling and predictive analysis	Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.