nature aging

ARTICLES

https://doi.org/10.1038/s43587-021-00044-4

Check for updates

Disparities in the pace of biological aging among midlife adults of the same chronological age have implications for future frailty risk and policy

Maxwell L. Elliott[1]✉, Avshalom Caspi[1,2,3,4,5], Renate M. Houts[1], Antony Ambler[3,6], Jonathan M. Broadbent[7], Robert J. Hancox[8], HonaLee Harrington[1], Sean Hogan[6], Ross Keenan[9,10], Annchen Knodt[1], Joan H. Leung[11,12], Tracy R. Melzer[9,13], Suzanne C. Purdy[9,11,12], Sandhya Ramrakha[6], Leah S. Richmond-Rakerd[14], Antoinette Righarts[8], Karen Sugden[1], W. Murray Thomson[7], Peter R. Thorne[9,12,15], Benjamin S. Williams[1], Graham Wilson[6], Ahmad R. Hariri[1], Richie Poulton[6] and Terrie E. Moffitt[1,2,3,4,5]

Some humans age faster than others. Variation in biological aging can be measured in midlife, but the implications of this variation are poorly understood. We tested associations between midlife biological aging and indicators of future frailty risk in the Dunedin cohort of 1,037 infants born the same year and followed to age 45. Participants' 'Pace of Aging' was quantified by tracking declining function in 19 biomarkers indexing the cardiovascular, metabolic, renal, immune, dental and pulmonary systems across ages 26, 32, 38 and 45 years. At age 45 in 2019, participants with faster Pace of Aging had more cognitive difficulties, signs of advanced brain aging, diminished sensory–motor functions, older appearances and more pessimistic perceptions of aging. People who are aging more rapidly than same-age peers in midlife may prematurely need supports to sustain independence that are usually reserved for older adults. Chronological age does not adequately identify need for such supports.

As we age, the risk that we will experience chronic diseases (for example, heart disease, diabetes and cancer) and declining capacities (for example, reduced strength, impaired hearing and poorer memory) increases[1]. To help mitigate personal and societal costs associated with aging, population-level policies typically specify eligibility on the basis of chronological age. These include retirement age, pensions, social security and healthcare subsidies, all intended to support independence. However, while many individuals continue to live independently and flourish into their nineties, others experience organ failure, dementia and mortality before their sixties, the age when entitlement to many of the aforementioned age-based supports begins[2]. Thus, chronological age is, at best, an imperfect basis for aging policy.

All individuals age chronologically at the same rate, but there is marked variation in their rate of biological aging; this may help explain why some adults experience age-related decline faster than others[3,4]. Biological aging can be defined as decline that (1) simultaneously involves multiple organ systems and (2) is gradual and progressive[5]. Across the lifespan, the consequences of individual differences in genetic endowment, cellular biology and life experiences accumulate, driving the divergence of biological age from chronological age for some people[6–9]. Among older adults of the same chronological age, those with accelerated biological aging (as measured by blood and DNA methylation biomarkers) are more likely to develop heart disease, diabetes and cancer and have a higher rate of cognitive decline, disability and mortality[10–16].

Current disease-management strategies usually treat and manage each age-related chronic disease independently[7]. In contrast, the geroscience hypothesis proposes that many age-related chronic diseases could be prevented by slowing biological aging itself[7,17]. The geroscience hypothesis states that biological aging drives cellular-level deterioration across all organ systems, thereby causing the exponential rise in multimorbidity across the second half of the lifespan[6]. The implication is that by slowing biological aging directly, instead of managing each disease separately, the risk for all chronic age-related diseases could be simultaneously ameliorated[5]. Early trials suggest that this goal may be attainable[18,19]. To achieve maximal prevention of age-related diseases, interventions to slow biological aging will need to target individuals by midlife before decades of subclinical organ decline have accumulated[6,20]. However, little is known about how to identify adults in midlife who are aging fast and who are most likely to benefit from

[1]Department of Psychology and Neuroscience, Duke University, Durham, NC, USA. [2]Center for Genomic and Computational Biology, Duke University, Durham, NC, USA. [3]Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [4]PROMENTA, Department of Psychology, University of Oslo, Oslo, Norway. [5]Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA. [6]Dunedin Multidisciplinary Health and Development Research Unit, Department of Psychology, University of Otago, Dunedin, New Zealand. [7]Faculty of Dentistry, University of Otago, Dunedin, New Zealand. [8]Department of Preventive and Social Medicine, Otago Medical School, University of Otago, Dunedin, New Zealand. [9]Brain Research New Zealand-Rangahau Roro Aotearoa, Centre of Research Excellence, Universities of Auckland and Otago, Auckland, New Zealand. [10]Christchurch Radiology group, Christchurch, New Zealand. [11]School of Psychology, University of Auckland, Auckland, New Zealand. [12]Eisdell Moore Centre, University of Auckland, Auckland, New Zealand. [13]Department of Medicine, University of Otago, Dunedin, New Zealand. [14]Department of Psychology, University of Michigan, Ann Arbor, MI, USA. [15]School of Population Health, University of Auckland, Auckland, New Zealand. ✉e-mail: maxwell.elliott@duke.edu

geroscience-informed interventions and, for this reason, we studied biological aging in midlife.

We measured biological aging in a population-representative 1972–1973 birth cohort of 1,037 individuals followed from birth to age 45 years in 2019 with 94% retention: the Dunedin Study[21]. Over 20 years—at ages 26, 32, 38 and 45—we repeatedly collected 19 biomarkers to assess changes in the function of cardiovascular, metabolic, renal, immune, dental and pulmonary systems, and quantified age-related decline shared among these systems (Fig. 1). We call this index of biological aging in the Dunedin Study the 'Pace of Aging'. In a 2015 article, we quantified the Pace of Aging across 12 years from age 26 to 32 to 38 (ref. [22]). Here, we report three innovations. First, we extend measurement to age 45, which yields a measure that exceeds the midpoint of the contemporary lifespan. Second, we report four biomarker waves, which increases statistical power for growth–curve models of biomarker-decline slopes and improves precision of the Pace of Aging. Third, we test associations with new outcomes: brain structure, brain age gap estimate (brainAGE), gait speed, additional function tests, visual contrast sensitivity, hearing and attitudes toward aging.

We tested the hypothesis that individual differences in the Pace of Aging from ages 26 to 45 would be associated, at age 45, with established risk factors for future frailty, morbidity and early mortality across four domains (Fig. 1). First, we tested whether individuals with a faster Pace of Aging had, at 45, early signs of brain aging that have been linked to dementia in older adults. Second, we tested whether individuals with a faster Pace of Aging had more cognitive difficulties and cognitive decline. Third, we tested whether those with a faster Pace of Aging already displayed signs of diminished sensory–motor functional capacities that are linked to loss of independence, falls and mortality in studies of older adults. Fourth, we tested whether individuals with an accelerated Pace of Aging look older than their same-aged peers, whether they self-report pessimism about aging and whether informants have noticed age-related difficulties in study members.

Results

Quantifying two decades of biological aging in midlife. The Pace of Aging was quantified in three steps. First, we measured longitudinal changes in 19 biomarkers at ages 26, 32, 38 and 45 years (see Supplementary Table 1 for details on each biomarker), assessing cardiovascular, metabolic, renal, immune, dental and pulmonary systems, totaling 69,715 data points (cohort participants × biomarkers × assessment phases) (Fig. 1). All biomarkers at each age were standardized on the basis of their original distribution at age 26 (that is, set to a mean of 0 and a standard deviation (s.d.) of 1) and coded so that higher values represented 'older/less healthy' levels (that is, scores were reversed for cardiovascular fitness, lung function, creatinine clearance and high-density lipoprotein cholesterol for which values are expected to decline with increasing chronological age). In our cohort of midlife adults, biomarkers showed a pattern of age-dependent decline in the functioning of multiple organ systems over the 20-year follow-up period.

Second, linear mixed-effects modeling was used to quantify each study member's personal rate of change across each of the 19 biomarkers. The 19 models took the form $B_{it}=\gamma_0+\gamma_1 \mathrm{Age}_{it}+\mu_{0i}+\mu_{1i}\mathrm{Age}_{it}+\epsilon_{it}$, where B_{it} is a biomarker measured for individual i at time t, γ_0 and γ_1 are the fixed intercept and slope estimated for the cohort, and μ_{0i} and μ_{1i} are the random intercepts and slopes estimated for each individual i. Biomarker slopes indicated a tendency to decline with age (Fig. 2a). Of the 171 unique correlations among biomarker slopes, 124 (73%) had a positive sign indicating coordinated change with age. Correlations between biomarker slopes averaged $r=0.1$ ranging from $r=-0.2$ to $r=0.7$ across the 19 biomarkers (Supplementary Table 2).

Third, we combined information from the 19 slopes to calculate each study member's personal Pace of Aging. In line with the geroscience hypothesis, which states that aging represents correlated gradual decline across organ systems, we calculated each study member's Pace of Aging as the sum of age-dependent annual changes across all biomarkers: Pace of Aging$_i$ = $\sum_{B=1}^{19}\mu_{1iB}$. The resulting Pace of Aging was then scaled to a mean of 1 so that it could be interpreted with reference to an average rate of 1 year of biological aging per year of chronological aging. Study members showed wide variation in their Pace of Aging (mean = 1 biological year per chronological year, s.d. = 0.29). Over the two decades in which we measured biological aging, the study member with the slowest Pace of Aging aged by just 0.40 biological years per chronological year, while the study member with the fastest Pace of Aging accrued 2.44 biological years per chronological year (Fig. 2b).

Accelerated biological aging and the aging brain. Deterioration of the brain (for example, in Alzheimer's disease and related dementias) is a major contributor to morbidity and loss of independence in older adults[23,24]. Brain imaging can detect subtle signs of brain aging decades before the onset of age-related disease[25,26]. Several magnetic resonance imaging (MRI) measures have been associated with a higher risk for cognitive decline and neurodegenerative disease in older adults, including thinner cortex, smaller surface area, smaller hippocampal volume, a larger volume of white matter hyperintensities (WMH), lower fractional anisotropy and older brain age[27–29]. Here, derived from high-resolution structural MRI scans at age 45, we found that an accelerated Pace of Aging in the first half of the lifespan was associated with most of these risk factors. Table 1 reports effect sizes, significance tests and covariate-adjusted analyses (sex-, continuously measured body mass index (BMI)- and smoking-adjusted; and excluding study members with cancer, diabetes or heart attack) for all brain measures.

Study members with a faster Pace of Aging had thinner average cortical thickness ($\beta=-0.14$, $P<0.001$; 95% confidence interval (CI): −0.21, −0.08) and smaller total surface area of the cortex ($\beta=-0.08$, $P=0.003$; 95% CI: −0.14, −0.03). Furthermore, regional investigations of cortical thickness and surface area were conducted using the HCP-MPP1.0 parcellation, which consists of 360 parcels (that is, brain regions)[30]. Associations between faster Pace of Aging and thinner cortex were widespread (89.72% of parcels had negative effect sizes, 38.33% were statistically significant at $P<0.05$, corrected for false discovery rate; Supplementary Table 3), with the largest associations in the medial temporal and insular cortices (Fig. 3a). Regional associations with surface area were also widespread (96.11% of parcels had negative effect sizes, 22.50% were statistically significant, $P<0.05$, corrected for false discovery rate; Supplementary Table 4), with the largest associations in the visual and lateral temporal cortices (Fig. 3b). Those with a faster Pace of Aging also had smaller volumes of the hippocampus ($\beta=-0.10$, $P=0.001$; 95% CI: −0.16, −0.04)—a brain region central to both healthy memory function and age-related memory decline[31].

Study members with a faster Pace of Aging had early signs of white matter deterioration, as indicated by a larger log-transformed volume of WMH ($\beta=0.18$, $P<0.001$; 95% CI: 0.11, 0.24; Fig. 3c), but they did not have lower fractional anisotropy ($\beta=-0.03$, $P=0.439$; 95% CI: −0.09, 0.04), a measure of white matter microstructural integrity.

We also studied a relatively new measure called 'brain age gap estimate' (brainAGE). BrainAGE is the difference between each study member's chronological age and their brain age as estimated from a machine-learning algorithm that was trained to predict chronological age from gray- and white matter measures in independent samples ranging in age from 19 to 82 (ref. [32]). Higher scores on brainAGE thus indicate a brain age that is older than chronological age. Dunedin study members with a faster Pace of Aging tended to have brains that were typical of an older person as represented by higher brainAGE scores ($\beta=0.20$, $P<0.001$; 95% CI: 0.13, 0.26; Fig. 3d).

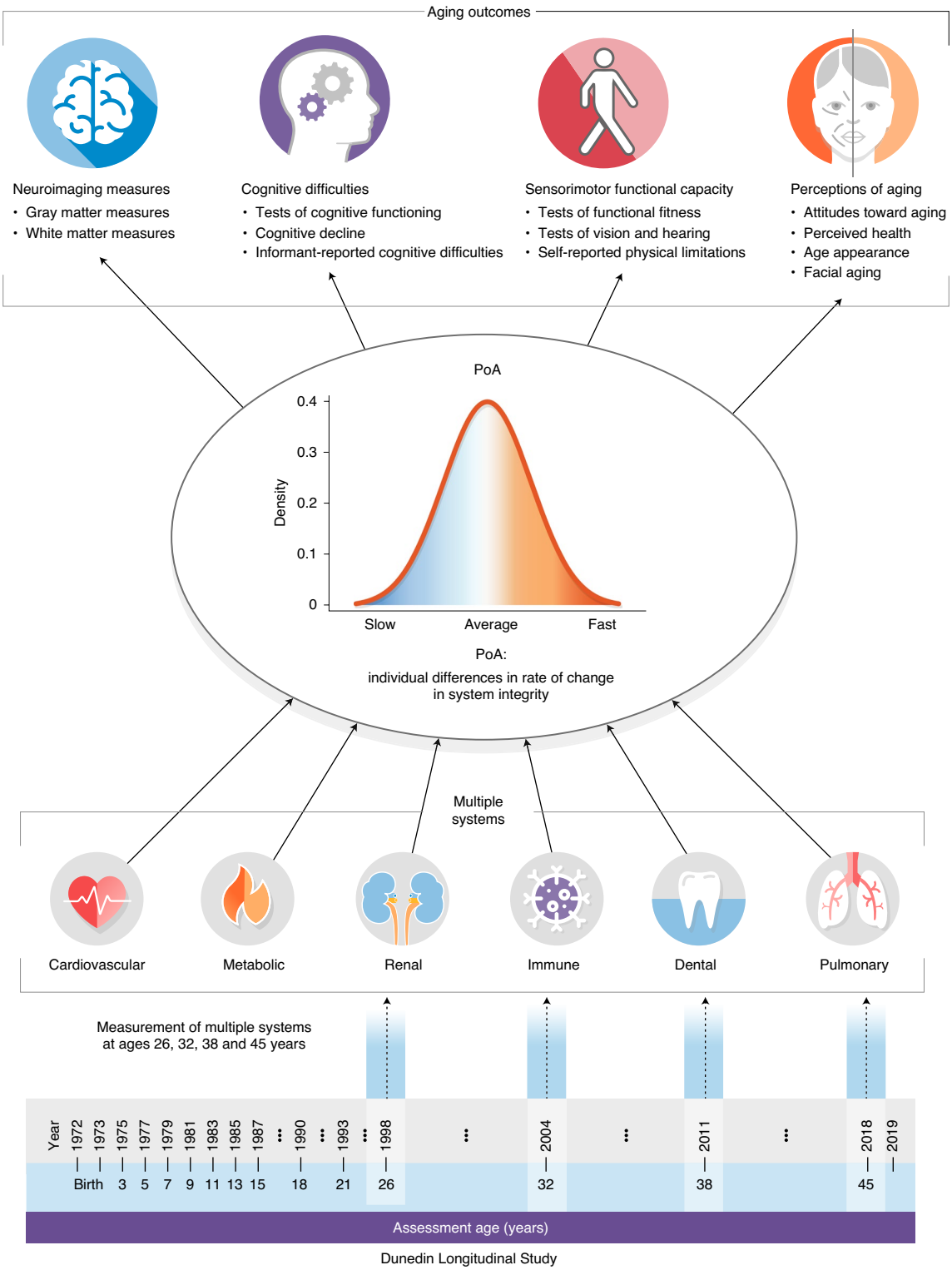

Fig. 1 | Study design. We studied the Pace of Aging (PoA) in the Dunedin birth cohort. The timeline on the bottom of the figure visualizes the design of the Dunedin Longitudinal Study. The years of each phase of data collection and the corresponding ages are listed. The PoA was derived from measuring longitudinal changes in 19 biomarkers at four timepoints between the ages of 26 and 45 years. These biomarkers indexed functioning across multiple organ systems (each visualized under the heading 'multiple systems'). We combined rates of changes across these biomarkers to produce a single measure termed the PoA. We then investigated associations between the PoA and aging outcomes across four domains at age 45: neuroimaging measures, cognitive difficulties, sensorimotor functional capacity and perceptions of aging.

Accelerated biological aging, cognitive difficulties and cognitive decline. Cognitive testing is used widely to assess risk for age-related neurological disease. Low cognitive functioning is a risk factor for Alzheimer's disease and dementia, and cognitive decline is a hallmark feature of these age-related disorders[33,34]. Dunedin study members with a faster Pace of Aging displayed poorer cognitive

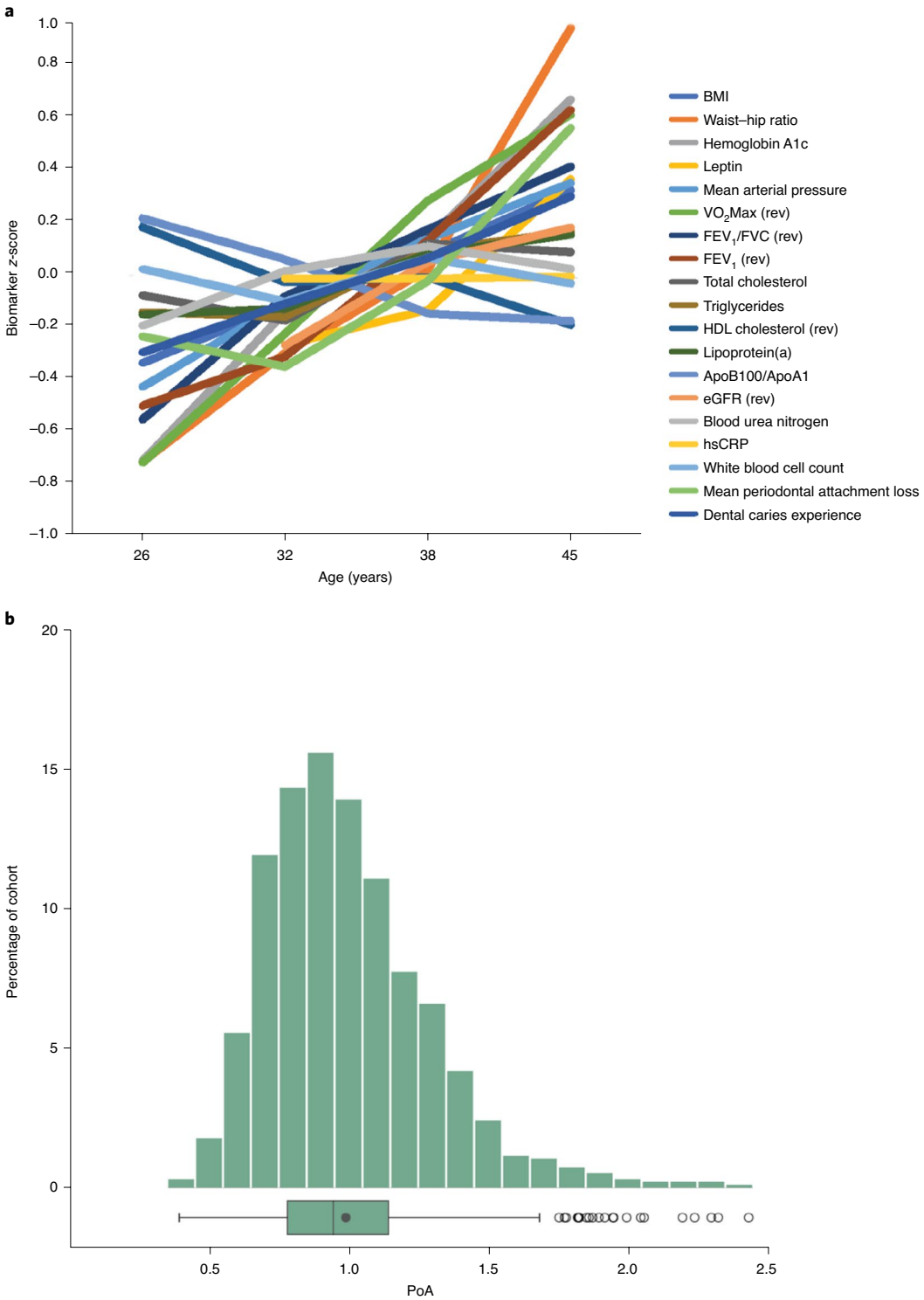

Fig. 2 | Biological aging across two decades from age 26 to age 45. a, For visualization, biomarker values were standardized to have mean = 0 and s.d. = 1 across the two decades of follow up (*z*-scores). *Z*-scores were coded so that higher values corresponded to older levels of the biomarkers. **b**, PoA is denominated in years of physiological change per chronological year. A PoA of 1 indicates a cohort member who experienced 1 year of physiological change per chronological year (the cohort average). A PoA of 2 indicates a cohort member aging at a rate of 2 years of physiological change per chronological year (that is, twice as fast as the cohort average). The box plot displays the distribution of the PoA; the box borders and midline represent the 25th, 50th and 75th percentiles, with whiskers extending to the furthest observation within the 1.5 interquartile range of the 25th and 75th percentiles. $n = 955$ study members. Apo, apolipoprotein; (rev), reverse.

functioning and more cognitive decline by age 45. Table 1 reports effect sizes, significance tests and covariate-adjusted analyses for all cognitive measures.

Compared to peers with a slower Pace of Aging, those who were aging faster had lower age 45 intelligence quotient (IQ) scores ($\beta = -0.33$, $P < 0.001$; 95% CI: −0.38, −0.26). This difference in

Table 1 | Associations between the Pace of Aging, neuroimaging and cognitive measures

	Adjusted for sex			Adjusted for sex, BMI and smoking			Without cancer, diabetes or heart attack		
	n	*β* (95% CI)	*P*	*n*	*β* (95% CI)	*P*	*n*	*β* (95% CI)	*P*
Signs of brain aging									
Cortical thickness	860	−0.14 (−0.21 to −0.08)	<0.001	858	−0.15 (−0.24 to −0.05)	0.002	806	−0.14 (−0.22 to −0.07)	<0.001
Surface area	860	−0.08 (−0.14 to −0.03)	0.003	858	−0.13 (−0.20 to −0.05)	0.001	806	−0.06 (−0.12 to −0.01)	0.031
Hippocampal volume	860	−0.10 (−0.16 to −0.04)	0.001	858	−0.14 (−0.22 to −0.06)	0.001	806	−0.08 (−0.15 to −0.02)	0.009
log[WMH volume]	851	0.18 (0.11 to 0.24)	<0.001	848	0.17 (0.07 to 0.26)	<0.001	797	0.18 (0.12 to 0.26)	<0.001
Fractional anisotropy	853	−0.03 (−0.09 to 0.04)	0.439	852	−0.12 (−0.22 to −0.03)	0.010	802	−0.02 (−0.09 to 0.05)	0.620
BrainAGE	868	0.20 (0.13 to 0.26)	<0.001	865	0.20 (0.11 to 0.29)	<0.001	813	0.18 (0.11 to 0.25)	<0.001
Cognitive difficulties									
Full-scale IQ	916	−0.33 (−0.38 to −0.26)	<0.001	910	−0.32 (−0.41 to −0.24)	<0.001	859	−0.33 (−0.40 to −0.27)	<0.001
IQ decline (residualized change)	904	−0.16 (−0.22 to −0.09)	<0.001	899	−0.18 (−0.27 to −0.09)	<0.001	847	−0.17 (−0.25 to −0.11)	<0.001
Verbal comprehension index	893	−0.30 (−0.36 to −0.24)	<0.001	889	−0.35 (−0.44 to −0.26)	<0.001	838	−0.31 (−0.39 to −0.25)	<0.001
Perceptual reasoning index	904	−0.27 (−0.33 to −0.20)	<0.001	900	−0.27 (−0.36 to −0.18)	<0.001	848	−0.26 (−0.34 to −0.20)	<0.001
Working memory index	900	−0.22 (−0.28 to −0.15)	<0.001	896	−0.18 (−0.27 to −0.09)	<0.001	844	−0.22 (−0.30 to −0.16)	<0.001
Processing speed index	904	−0.23 (−0.29 to −0.16)	<0.001	900	−0.24 (−0.33 to −0.15)	<0.001	848	−0.22 (−0.29 to −0.16)	<0.001
RAVL learning memory	905	−0.29 (−0.34 to −0.22)	<0.001	901	−0.28 (−0.37 to −0.20)	<0.001	849	−0.21 (−0.28 to −0.15)	<0.001
RAVL recall	901	−0.19 (−0.25 to −0.13)	<0.001	897	−0.20 (−0.29 to −0.12)	<0.001	845	−0.30 (−0.37 to −0.25)	<0.001
Informant memory difficulties	881	0.15 (0.08 to 0.21)	<0.001	875	0.24 (0.15 to 0.33)	<0.001	827	0.16 (0.10 to 0.24)	<0.001
Informant attention difficulties	881	0.20 (0.14 to 0.26)	<0.001	875	0.29 (0.20 to 0.38)	<0.001	827	0.22 (0.16 to 0.29)	<0.001

On the left side of the table are associations from the main text from linear regression models that were adjusted for sex. In the middle are sensitivity analyses in which the models also adjusted for age 45, BMI and smoking. On the right are the results from models that are adjusted for sex in which all study members who had a diagnosis of cancer, diabetes or heart attack by age 45 were excluded ($n = 58$). The outcome variables are grouped into modalities labeled in bold. All statistically significant (two-sided) sex-adjusted associations remain statistically significant after false discovery rate correction for the 38 tests presented in Tables 1 and 2.

cognitive functioning reflected actual cognitive decline over the years: when we compared age 45 IQ test scores with baseline scores from the childhood version on the same IQ test, study members with a faster Pace of Aging tended to show decline net of their baseline level ($\beta = -0.16$, $P < 0.001$; 95% CI: −0.22, −0.09). Furthermore, a faster Pace of Aging was associated broadly with poorer cognitive functioning across domains: study members with a faster Pace of Aging had poorer verbal comprehension ($\beta = -0.30$, $P < 0.001$; 95% CI: −0.36, −0.24), perceptual reasoning ($\beta = -0.27$, $P < 0.001$; 95% CI: −0.33, −0.20), working memory ($\beta = -0.22$, $P < 0.001$; 95% CI: −0.28, −0.15), processing speed ($\beta = -0.23$, $P < 0.001$; 95% CI: −0.29, −0.16), worse memory learning performance (Rey auditory verbal learning (RAVL): learning memory, $\beta = -0.29$, $P < 0.001$; 95% CI: −0.34, −0.22) and worse delayed memory recall (RAVL: recall, $\beta = -0.19$, $P < 0.001$; 95% CI: −0.25, −0.13).

Cognitive difficulties were detectable not only in objective tests but also noticeable in everyday life. Informants, who were surveyed because they knew a study member well, reported that study members with a faster Pace of Aging experienced more memory difficulties ($\beta = 0.15$, $P < 0.001$; 95% CI: 0.08, 0.21) and attention problems ($\beta = 0.20$, $P < 0.001$; 95% CI: 0.14, 0.26); for example, they noted that faster-aging study members were more likely to be 'easily distracted' and 'get sidetracked' as well as to 'misplace wallet, keys or eyeglasses' and 'forget to do errands, return calls or pay bills'.

Accelerated biological aging and diminished sensory–motor functional capacities. In gerontology, poor scores on tests of sensory–motor functioning (for example, gait speed, grip strength, visual contrast sensitivity and hearing thresholds) are often used to identify frail individuals who are at high risk for falls, loss of independence and mortality[35–38]. Dunedin study members who were aging faster showed several signs of sensory–motor difficulties. Table 2 reports effect sizes, significance tests and covariate-adjusted analyses for all sensory–motor measures.

Table 2 | Associations between the Pace of Aging, measures of sensory–motor functional capacity and perceptions of aging

	Adjusted for sex			Adjusted for sex, BMI and smoking			Without cancer, diabetes or heart attack		
	n	*β* (95% CI)	*P*	*n*	*β* (95% CI)	*P*	*n*	*β* (95% CI)	*P*
Sensory–motor capacity									
Gait speed	903	−0.33 (−0.39 to −0.27)	<0.001	901	−0.19 (−0.27 to −0.10)	<0.001	848	−0.33 (−0.40 to −0.27)	<0.001
One-legged balance	909	−0.36 (−0.42 to −0.30)	<0.001	905	−0.26 (−0.34 to −0.17)	<0.001	853	−0.35 (−0.42 to −0.29)	<0.001
Chair stands	872	−0.30 (−0.37 to −0.24)	<0.001	871	−0.25 (−0.34 to −0.16)	<0.001	820	−0.30 (−0.38 to −0.24)	<0.001
Step in place	885	−0.28 (−0.34 to −0.22)	<0.001	884	−0.21 (−0.30 to −0.12)	<0.001	832	−0.28 (−0.36 to −0.22)	<0.001
Grip strength	919	−0.05 (−0.09 to −0.01)	0.017	913	−0.10 (−0.15 to −0.04)	<0.001	863	−0.03 (−0.07 to 0.01)	0.110
Grooved pegboard	901	−0.27 (−0.33 to −0.20)	<0.001	897	−0.21 (−0.29 to −0.12)	<0.001	845	−0.27 (−0.34 to −0.21)	<0.001
Contrast sensitivity	903	−0.13 (−0.19 to −0.06)	<0.001	899	−0.10 (−0.19 to −0.01)	0.036	847	−0.11 (−0.18 to −0.04)	0.002
Audiometry: HF-PTA	900	0.17 (0.10 to 0.23)	<0.001	897	0.14 (0.05 to 0.23)	0.003	845	0.19 (0.13 to 0.26)	<0.001
Audiometry: 4F-PTA	901	0.20 (0.14 to 0.26)	<0.001	897	0.14 (0.05 to 0.23)	0.003	846	0.14 (0.07 to 0.21)	<0.001
LiSN-S low cue	901	−0.17 (−0.23 to −0.10)	<0.001	897	−0.10 (−0.19 to −0.01)	0.026	845	−0.17 (−0.24 to −0.10)	<0.001
LiSN-S spatial advantage	901	0.22 (0.15 to 0.28)	<0.001	897	0.18 (0.09 to 0.27)	<0.001	845	0.21 (0.14 to 0.21)	<0.001
Physical limitations (SF-36)	921	0.29 (0.23 to 0.35)	<0.001	913	0.11 (0.02 to 0.19)	0.012	864	0.29 (0.23 to 0.36)	<0.001
Perceptions of aging									
Self-reported aging attitudes	920	−0.22 (−0.28 to −0.16)	<0.001	911	−0.25 (−0.34 to −0.16)	<0.001	863	−0.21 (−0.28 to −0.14)	<0.001
Self-reported health	927	−0.35 (−0.41 to −0.29)	<0.001	916	−0.27 (−0.36 to −0.19)	<0.001	870	−0.34 (−0.42 to −0.29)	<0.001
Self-reported perceived age	892	0.09 (0.03 to 0.16)	0.005	888	0.11 (0.01 to 0.20)	0.024	838	0.08 (0.02 to 0.16)	0.018
Informant-reported health	881	−0.38 (−0.45 to −0.32)	<0.001	875	−0.30 (−0.38 to −0.21)	<0.001	827	−0.36 (−0.44 to −0.31)	<0.001
Researcher-reported health	930	−0.58 (−0.62 to −0.52)	<0.001	916	−0.45 (−0.51 to −0.37)	<0.001	873	−0.56 (−0.62 to −0.51)	<0.001
Informant-reported age appearance	881	0.35 (0.29 to 0.41)	<0.001	875	0.34 (0.25 to 0.43)	<0.001	827	0.34 (0.29 to 0.42)	<0.001
Researcher-reported age appearance	930	0.44 (0.38 to 0.49)	<0.001	916	0.40 (0.31 to 0.47)	<0.001	873	0.43 (0.37 to 0.50)	<0.001
Self-reported age appearance	894	0.10 (0.03 to 0.16)	0.003	891	0.11 (0.02 to 0.21)	0.015	838	0.08 (0.01 to 0.15)	0.029
Facial age	905	0.33 (0.26 to 0.39)	<0.001	901	0.30 (0.22 to 0.39)	<0.001	850	0.33 (0.28 to 0.41)	<0.001
Perceived longevity	908	−0.27 (−0.33 to −0.20)	<0.001	902	−0.20 (−0.28 to 0.11)	<0.001	852	−0.26 (−0.32 to −0.19)	<0.001

On the left side of the table are associations from the main text from linear regression models that were adjusted for sex. On the right side of the table are sensitivity analyses in which the models also adjusted for age 45 BMI and smoking. On the right are the results from models that are adjusted for sex in which all study members who had a diagnosis of cancer, diabetes or heart attack by age 45 were excluded ($n = 58$). The outcome variables are grouped into modalities labeled in bold. All statistically significant (two-sided) associations in this table remain statistically significant after false discovery rate correction for the 38 tests presented in Tables 1 and 2.

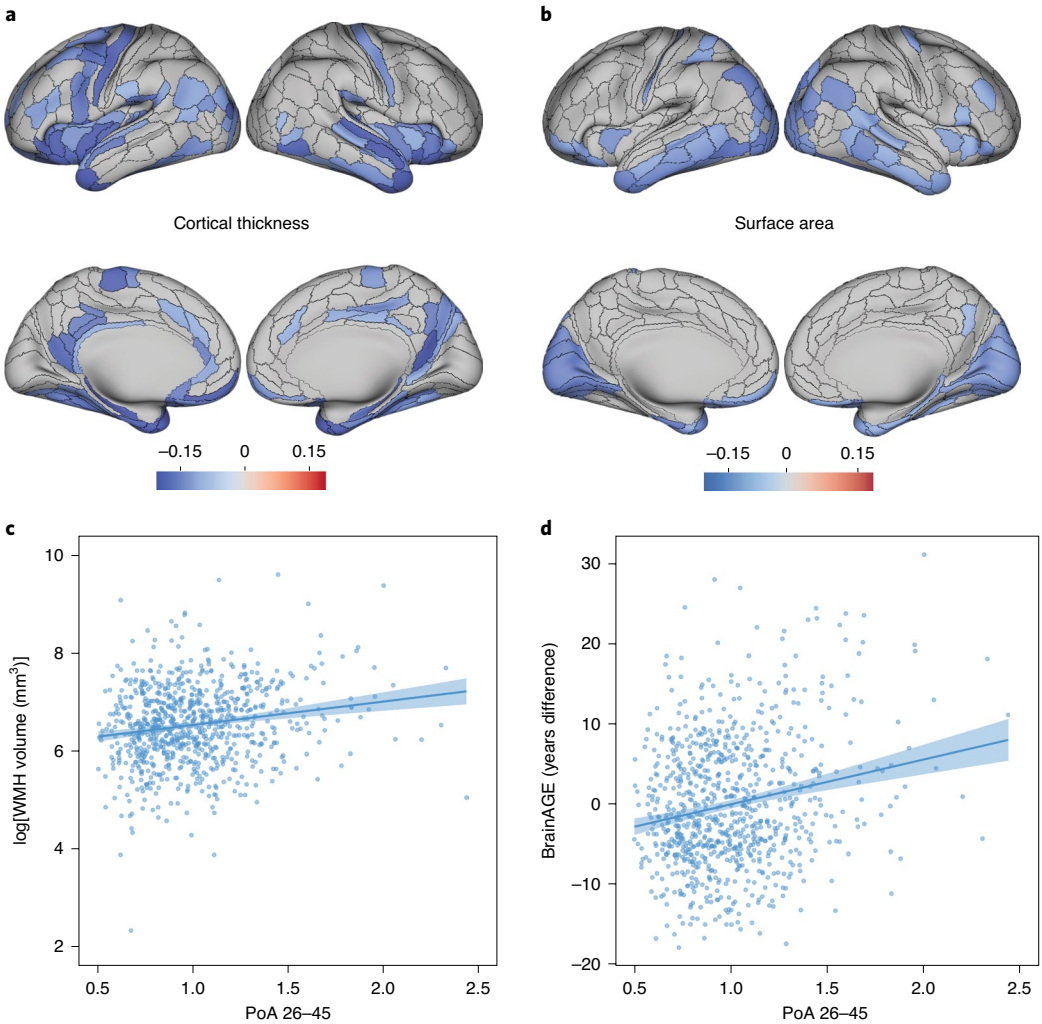

Fig. 3 | Study members who were aging faster showed signs of advanced brain aging relative to slower-aging peers. a,**b**, The overlays display cortical regions (in blue) whose thickness (**a**) or surface area (**b**) are significantly associated (false discovery rate corrected, two-sided test) with PoA. Associations were tested using linear regression performed at each cortical region. **c**,**d**, The scatterplots show associations between PoA and volume of WMH ($n=851$) (**c**) as well as brainAGE (a measure of the difference between each study member's chronological age and their brain age as estimated from a machine-learning algorithm that was trained to predict chronological age from gray- and white matter measures in independent samples ranging in age from 19 to 82; $n=868$) (**d**)[32]. Scatterplots include the mean regression line ±1 s.e.m.

Compared to peers with a slower Pace of Aging, those who were aging faster had slower gait speed ($\beta=-0.33$, $P<0.001$; 95% CI: −0.39, −0.27), poorer balance (one-legged balance, $\beta=-0.36$, $P<0.001$; 95% CI: −0.42, −0.30), were slower at rising repeatedly from a chair (chair stands, $\beta=-0.30$, $P<0.001$; 95% CI: −0.37, −0.24) and stepping in place (2-min step test, $\beta=-0.28$, $P<0.001$; 95% CI: −0.34, −0.22), were weaker (grip strength, $\beta=-0.05$, $P=0.017$; 95% CI: −0.09, −0.01) and had more difficulties with fine motor control (grooved pegboard, $\beta=-0.27$, $P<0.001$; 95% CI: −0.33, −0.20).

In addition, study members who were aging faster had diminished sensory abilities. Visual contrast sensitivity and hearing ability are known to decline with advanced age[35,39]. Study members with a faster Pace of Aging at age 45 had more difficulty visually distinguishing an object from its background on tests of contrast sensitivity ($\beta=-0.13$, $P<0.001$; 95% CI: −0.19, −0.07). They also had more difficulties detecting high-pitch-tones (high-frequency pure-tone audiometry (HF-PTA), $\beta=0.17$, $P<0.001$; 95% CI: 0.10, 0.23) and low- to mid-pitch tones (four-frequency pure-tone average (4F-PTA), $\beta=0.20$, $P<0.001$; 95% CI: 0.14, 0.26) and were worse at hearing sentences in noisy environments when auditory distractors were nearby (listening in spatialised noise–sentences test (LiSN-S) low cue, $\beta=-0.17$, $P<0.001$; 95% CI: −0.23, −0.10) and when distractors were spatially distant (LiSN-S spatial advantage, $\beta=0.22$, $P<0.001$; 95% CI: 0.15, 0.28). Finally, study members with a faster Pace of Aging noticed sensory–motor difficulties in their everyday lives, self-reporting more physical limitations (36-item short-form health survey (SF-36) physical functioning scale, $\beta=0.29$, $P<0.001$; 95% CI: 0.23, 0.35).

Accelerated biological aging and negative perceptions of aging. Age-related morbidity and mortality are not only forecasted by objective measures of physical and cognitive functioning. Older adults who self-report that they feel old are also more likely to subsequently be diagnosed with age-related disease and die at a younger age[40,41]. We found that study members with a faster Pace of Aging were more likely to hold unfavorable views of aging. Table 2 reports effect sizes, significance tests and covariate-adjusted analyses for all perception measures.

Study members with a faster Pace of Aging had more negative attitudes towards aging ($\beta=-0.22$, $P<0.001$; 95% CI: −0.28, −0.16),

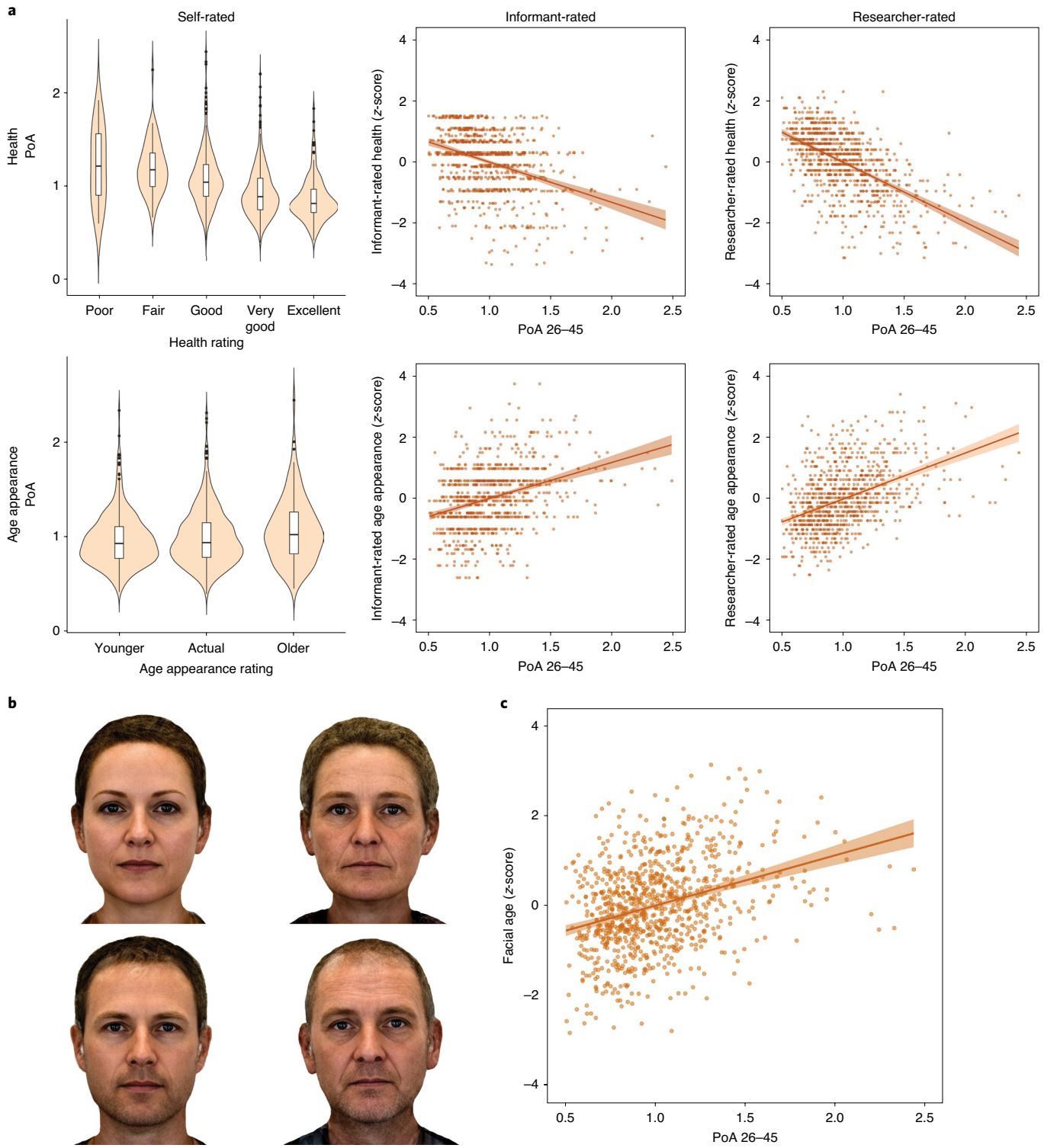

Fig. 4 | Study members who were aging faster were perceived as less healthy and looking older when compared to slower-aging peers. a, Associations between PoA and self-reported health ($n = 927$) and age appearance ($n = 892$), informant-rated health ($n = 881$) and age appearance ($n = 881$) and research-worker-rated health ($n = 930$) and age appearance ($n = 930$). Violin/box plots show the distribution of PoA at each self-rating; the box borders and midline represent the 25th, 50th and 75th percentiles, with whiskers extending to the furthest observation within 1.5 interquartile ranges of the 25th and 75th percentiles. **b**, Digitally averaged composite faces made up of the ten male and female study members with the youngest (left) and oldest (right) facial age ratings. **c**, Scatterplot of the association between PoA and facial age ratings by independent raters ($n = 905$). Scatterplots include the mean regression line ±1 s.e.m. All graphs are adjusted for sex.

endorsing sentiments such as 'things keep getting worse as I get older' and 'I am not as happy now as I was when I was younger'. They self-reported that they felt less healthy ($\beta = -0.35$, $P < 0.001$; 95% CI: −0.41, −0.29) and that they felt older than their chronological age ($\beta = 0.09$, $P = 0.005$; 95% CI: 0.03, 0.16). When asked similar questions about the study members, informants (who knew them well)

and research workers (who met the study members during their 1-day unit visit) both reported that study members with a faster Pace of Aging seemed in worse health (informant, $\beta = -0.38$, $P < 0.001$; 95% CI: −0.45, −0.32; research worker, $\beta = -0.58$, $P < 0.001$; 95% CI: −0.62, −0.52) and looked older than their age (informant, $\beta = 0.35$, $P < 0.001$; 95% CI: 0.29, 0.41; research worker, $\beta = 0.44$, $P < 0.001$; 95% CI: 0.38, 0.49; Fig. 4a). In addition, study members who were aging faster self-reported that they looked older than their age ($\beta = 0.10$, $P = 0.003$; 95% CI: 0.03, 0.16) and, when solely presented with facial images, independent raters scored study members with a faster Pace of Aging as looking older than their peers ($\beta = 0.33$, $P < 0.001$; 95% CI: 0.26, 0.39) (Fig. 4b,c). Finally, study members with a faster Pace of Aging were less likely to think that they would live past the age of 75 ($\beta = -0.27$, $P < 0.001$; 95% CI: −0.33, −0.20).

Sensitivity analyses. Associations with the Pace of Aging were robust to several sensitivity analyses. First, associations with the Pace of Aging were not attributable to being overweight or to smoking at age 45 (Tables 1 and 2). Second, associations with the Pace of Aging were not driven by common age-related diseases (Tables 1 and 2). Third, associations were not dependent on the slope of any single biomarker or any family of biomarkers included in the measure of the Pace of Aging (Supplementary Figs. 1 and 2). Fourth, associations were not driven by outlier values or by social class (Supplementary Fig. 3). Fifth, associations with the Pace of Aging were best characterized as linear (Supplementary Table 5). Sixth, associations with the Pace of Aging were not driven by baseline differences in health at age 26 (intercept; Supplementary Table 6). Finally, associations with the Pace of Aging were approximately equal in males and females (Supplementary Fig. 4 and Supplementary Table 7).

Discussion

Chronological age is a poor proxy for biological age, even in midlife. Here, in a population-representative birth cohort without variation in chronological age, we found that study members varied widely in their Pace of Aging. Furthermore, study members who had a faster Pace of Aging in midlife exhibited signs of advanced brain aging, experienced more cognitive difficulties, had diminished sensory–motor functional capacity and had more negative perceptions of aging. Sensitivity analyses supported our hypothesis that the Pace of Aging is a robust indicator of the cumulative, progressive and gradual deterioration across organ systems that underlies biological aging. Together, these findings support at least two conclusions: (1) meaningful variation in biological aging can be measured in midlife and (2) people with a faster rate of biological aging across the first half of the lifespan are more likely to experience age-related functional impairment by midlife. These findings reinforce the question of whether midlife is a window of opportunity for the mitigation of age-related disease[42–44]. We have shown that biological aging in midlife is meaningful, yet further research is needed to determine whether biological aging in midlife is still malleable. Randomized trials are beginning to test this possibility[45,46].

Four design features of the Dunedin Study support these conclusions. First, all study members were born in 1972–1973, which allows the direct measurement of individual differences in biological aging uncoupled from age and cohort effects[47,48]. Second, the Dunedin Study has very low attrition rates; unlike many longitudinal studies of older adults that have selective attrition and mortality, the full range of health is represented[49,50]. Third, the Dunedin Study has collected four waves of biological measurements from age 26 to age 45—a unique dataset allowing for more accurate estimates of biological aging. Fourth, although age-related diseases are uncommon in midlife, study members were assessed at age 45 with a battery of established measures that are commonly used in geriatric settings to predict frailty, morbidity and mortality.

This study was not without limitations. First, these findings are on the basis of a single birth cohort from New Zealand. Second, our study currently lacks follow up past the age of 45. Further investigation of the Pace of Aging in diverse cohorts and older adults is needed. Third, the Pace of Aging was derived from 19 biomarkers repeatedly assessed across 20 years, which will be infeasible for most studies of biological aging. However, we recently reported that a proxy for the Pace of Aging can now be quantified from genome-wide DNA methylation data extracted from a single cross-sectional blood draw[51]. This advance makes it possible for studies lacking four waves of biomarkers to incorporate the measure and extend this work; for example, it predicts disease and mortality in US and UK samples[51]. Fourth, while associations were consistent across domains and measures, effect sizes were generally moderate. However, these moderate associations between the Pace of Aging and midlife function probably reflect the cumulative effects of the aging process. Therefore, if the Pace of Aging truly measures the underlying aging process, the associations reported here should grow larger over time, as fast and slow agers continue to diverge[52].

Within the bounds of these limitations, our findings have implications for geroscience theory, research and policy. Concerning theory, Pace of Aging operationalises geroscience theory, unlike previous methods for measuring biological aging. Many biological-age measures are derived from cross-sectional multi-age datasets and therefore confound age with cohort differences[17,47,53]. People born 70–80 years ago and people born 20–30 years ago experienced differential exposures to childhood diseases, tobacco smoke, airborne lead, antibiotics, anti-inflammatory medications and poorer nutrition. Cross-sectional biological-age measures can also be confounded by acute illness that is not aging. In contrast, the Pace of Aging has four key features that discriminate age-related biological decline during adulthood (desired signal) versus health difficulties arising in early life from biomarker-altering exposures (noise), or acute illness near the time of blood draw (noise). Pace of Aging reflects (1) simultaneous decline in several organ systems, (2) decline in one direction that does not show recovery that would indicate acute illness, (3) decline that continues over 20 years and (4) decline in people all born the same year, unconfounded by cohort effects.

We find that a faster rate of biological aging, assessed across several organ systems, throughout midlife, is associated with several measures of functional impairment and frailty risk that have established links to morbidity and mortality in older adults. The breadth of these associations is consistent with the geroscience hypothesis depicting accelerated aging as a common cause of age-related chronic disease. Further research is needed to test whether interventions in humans can slow biological aging in midlife and reduce long-term risk for age-related chronic disease. Interventions that can achieve even mild slowing of biological aging promise to improve quality of life in older adults while yielding substantial healthcare savings[17,54].

Concerning research and policy, current efforts aimed at improving biological aging measurements are driven primarily by the need to test emerging anti-aging biotechnology[46,55,56]. While our findings support these efforts, they also suggest that biological aging may have broader implications for society. Many social programs, including state pensions and Medicare in the United States, are designed to offset the economic and health burdens that accrue as individuals age. Eligibility for these benefits has been determined historically on the basis of chronological age. For example, the age for US Social Security eligibility was set to 65 in 1939 when the average life expectancy was 63.7 (ref. [57]). However, with lengthening lifespans, it is important to also consider biological age. Our findings suggest that already by midlife, chronological age is a crude, poorly calibrated measure of the functional consequences of aging. We provide evidence that disparities in biological aging independent of chronological age are already linked to functional difficulties in midlife. Furthermore, by linking the Pace of Aging to

both objective and subjective outcomes, we found that the Pace of Aging carries a phenotypic presentation with biological and social implications. For example, study members who were in the fastest quintile of the Pace of Aging had brainAGEs that were an average of 3.79 years older and had faces that were rated as looking 4.32 years older than those in the slowest quintile. Thus, at age 45, study members with an accelerated Pace of Aging are simultaneously at higher risk for health challenges and future frailty as well as age-based discrimination.

Widespread application of biological aging measures could represent an alternative to using birthdates when determining the allocation of healthcare and financial support for those suffering from the sequelae of aging. For example, in the United States, there are continuing debates about lowering the Medicare age to expand access to preventative healthcare[58,59]. Perhaps someday we will be able to use biological aging measures to guide treatment access. With further development, geroscience could provide the conceptual tools, measurement technology and interventions required to mitigate disparities in the pace of biological aging through more tailored and just access to independence-sustaining resources.

Methods

Study design and population. Participants are members of the Dunedin Study, a longitudinal investigation of health and behavior in a representative birth cohort. The 1037 participants (91% of eligible births) were all individuals born between April 1972 and March 1973 in Dunedin, New Zealand, who were eligible on the basis of residence in the province and who participated in the first assessment at age 3 years[21]. The cohort represents the full range of socioeconomic status (SES) in the general population of New Zealand's South Island and, as adults, matches the New Zealand National Health and Nutrition Survey on key adult health indicators (for example, BMI, smoking and general practitioner visits) and the New Zealand Census of citizens of the same age on educational attainment[21,60]. The cohort is primarily white (93%, self-identified), matching South Island demographic characteristics. General assessments were performed at birth as well as ages 3, 5, 7, 9, 11, 13, 15, 18, 21, 26, 32 and 38 years and, most recently (completed April 2019), at age 45 years, when 938 of the 997 study members (94.1%) still alive participated. At the age 26-, 32-, 38- and 45-year assessments, biomarkers were collected that make up the Pace of Aging. Study members with data available at age 45 years did not differ significantly from other living participants in terms of childhood SES or childhood neurocognitive functioning (see attrition analysis in Supplementary Figs. 1 and 2). At each assessment, each participant was brought to the research unit for interviews and examinations. The research staff makes standardized ratings, informant questionnaires are collected and administrative records are searched. Written informed consent was obtained from cohort participants and study protocols were approved by the institutional ethical review boards of the participating universities. This study follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. The premise and analysis plan for this project were preregistered at https://bit.ly/2ZVtnsq.

Statistical analysis. Unless otherwise specified, all statistical analyses were completed using linear regression models in R (v.3.4.0). Unless otherwise noted, standardized regression coefficients are reported as Pearson's *r* effect sizes. All models were adjusted for sex. In addition, eight types of sensitivity analysis were performed to determine the robustness of the associations with the Pace of Aging. (1) In addition to sex, BMI and smoking status at age 45 were simultaneously added as covariates to rule out the possibility that associations were limited to overweight individuals and smokers (Tables 1 and 2). (2) Sex-adjusted models were run in which all study members who had diagnosed, common age-related diseases (cancer, diabetes or heart attack; Tables 1 and 2) were excluded. (3) To test the possibility that associations with the Pace of Aging were driven by a particular biomarker or family of biomarkers, we investigated all associations after systematically leaving out each of the 19 biomarkers in the Pace of Aging one at a time (Supplementary Fig. 1) and then after leaving out each family of biomarkers one at a time (Supplementary Fig. 2). (4) To test the possibility that associations with the Pace of Aging were biased by the long right-hand tail of the Pace of Aging distribution, we investigated associations after Winsorizing (±2 s.d.) the Pace of Aging and after log-transforming the Pace of Aging (Supplementary Fig. 3). (5) To test the possible influence of social determinants on aging trajectories as confounding factors, childhood SES was added as a covariate (Supplementary Fig. 3). (6) To test the possibility that associations with the Pace of Aging were better characterized as nonlinear associations, linear and quadratic Pace of Aging terms were added simultaneously as independent variables into the regression models (Supplementary Table 5). (7) To test the possibility that associations with the Pace of Aging were driven by baseline differences in health levels, the intercept from the Pace of Aging was entered as a covariate (Supplementary Table 6). (8) Finally, associations with the Pace of Aging were investigated separately for males and females to investigate the possibility that there were sex-differences in these associations (Supplementary Fig. 4). Multiple comparisons were corrected using the false discovery rate correction across all 38 sex-adjusted models presented in Tables 1 and 2. Analyses reported here were checked for reproducibility by an independent data-analyst who recreated the code by working from the manuscript and applied it to an independently generated copy of the dataset.

Measuring the Pace of Aging. We had four major goals in developing the measure of the Pace of Aging. (1) We sought to create a longitudinal biomarker panel that would capture the age-related decline of several different organ systems (for example, pulmonary, renal, dental, etc.). This is essential, as the idea of measuring the Pace of Aging is derived from geroscience, a field that—in contrast to single-disease paradigms—aims to understand how mechanisms of aging underlie multiple and diverse age-related diseases. (2) For a biomarker to be included in the Pace of Aging, we required evidence that each biomarker had been robustly associated with an age-related disease or early death in previous research. (3) We needed each biomarker to have a minimum of three waves of data to model the rate of decline in each biomarker using growth–curve modeling. In the Dunedin Study, this meant that we had to have data going back nearly 20 years to the late 1990s when we started our biobank (16 biomarkers have four waves and 3 had three waves). (4) Each biomarker had to be widely and routinely used so that our findings would be both translatable to clinical settings and generalizable to other studies. The Pace of Aging consists of all biomarkers that met these criteria.

Applying these criteria to the Dunedin Study biobank generated our panel of 19 biomarkers: BMI, waist–hip ratio, glycated hemoglobin, leptin, blood pressure (mean arterial pressure), cardiorespiratory fitness (maximal aerobic capacity VO_2Max), forced vital capacity ratio (FEV_1/FVC, where FVC is forced vital capacity), forced expiratory volume in 1 s (FEV_1), total cholesterol, triglycerides, high-density lipoprotein (HDL), lipoprotein(a), apolipoprotein B100/A1 ratio, estimated glomerular filtration rate (eGFR), blood urea nitrogen (BUN), high-sensitivity C-reactive protein (hsCRP), white blood cell count, mean periodontal attachment loss (AL) and the number of dental-caries-affected tooth surfaces (tooth decay). Biomarkers were assayed at the age 26, 32, 38 and 45 assessments. The Pace of Aging reported here represents an extension of a previously reported measure that used 18 biomarkers assayed at ages 26, 32 and 38 (ref. [22]). Here we add a recently completed fourth measurement wave of data, at age 45, totaling 19 biomarkers. We added measures of leptin and caries-affected tooth surfaces, both of which have now been assessed at multiple waves allowing growth–curve modeling. Telomere length was dropped because of an emerging and yet-unresolved field-wide debate about its measurement[61]. Specifically, telomere length derived from qPCR has been determined to be unsuitable for use in large epidemiological studies because of high levels of measurement error. Details on biomarker measurements are provided in Supplementary Table 1.

We calculated each study member's Pace of Aging in three steps. In the first step, we transformed the biomarker values to a standardized scale. For each biomarker at each wave, we standardized values according to the age 26 distribution (that is set to mean of 0 and s.d. of 1). Standardization was conducted separately for men and women. Standardized biomarker values greater than zero indicated levels that were 'older' and values less than zero indicated levels 'younger' than the average 26-year-old. To match, scores were reversed for VO_2Max, FEV_1/FVC, FEV_1, eGFR and HDL cholesterol, which are known to decline with age. Over the two decades of follow up, the biomarkers in the panel indicated a progressive deterioration of physiological integrity with advancing chronological age; that is, their cohort mean values tended to increase (that is, worsen) from the age 26 assessment to the age 45 assessment (Fig. 2).

In the second step, we calculated each study member's slope for each of the 19 biomarkers—the average year-on-year change observed over the two-decade period. Slopes were estimated using a mixed-effects growth model that regressed the biomarker's level on age. A complete list of means of biomarker slopes and pairwise correlations among biomarker slopes is presented in Supplementary Table 2. For only 4 of the 19 biomarkers we examined, cohort mean levels did not worsen over time, as expected on the basis of published associations with age-related chronic disease: white blood cell count and CRP levels remained stable with age; HDL cholesterol and apolipoprotein B100/A1 ratio improved with age. However, individual–difference slopes for these biomarkers did show the expected pattern of correlation with other biomarkers' slopes. For example, study members whose apolipoprotein B100/A1 ratio increased during the follow-up period also showed increasing adiposity, declining lung function and increasing systemic inflammation. We retained all preregistered biomarkers in the Pace of Aging model.

In the third step, we combined information from the 19 slopes of the biomarkers to calculate each study member's personal 'Pace of Aging'. Because we did not have any a priori basis for weighting differential contributions of the biomarkers to an overall Pace of Aging measure, we combined information using a unit-weighting scheme (all biomarkers were standardized to have mean = 0, s.d. = 1 on the basis of their age 26 distributions, so slopes were denominated in comparable units). We calculated each study member's Pace of Aging as the sum of

age-dependent annual changes in biomarker *Z*-scores. Because the Dunedin birth cohort represents its population, its mean and distribution represent population norms. We used these norms to scale the Pace of Aging to reflect physiological change relative to the passage of time. We set the cohort mean Pace of Aging as a reference value equivalent to the physiological change expected during a single chronological year. Using this reference value, we rescaled Pace of Aging in terms of years of physiological change per chronological year (mean = 1, s.d. = 0.29). On this scale, cohort members ranged in their Pace of Aging from 0.4 years of physiological change per chronological year (slow) to 2.4 years of physiological change per chronological year (fast) (Fig. 2).

As a sensitivity check to ensure that the geroscience definition of aging as unidirectional decline fits the data, we examined biomarker patterns of change for potential nonlinearity. Three biomarkers—leptin, hsCRP and eGFR—were measured at only three timepoints and could only be fit with a linear model. For all other biomarkers, we fit an additional model that included fixed effects for the intercept, linear change and quadratic change, as well as random effects for the intercept and linear terms. For nine biomarkers, fit statistics (residual log likelihood, Akaike information criterion and Bayesian information criterion) indicated that the linear model provided a better fit than the quadratic model. For seven biomarkers, fit statistics indicated that the quadratic model provided a marginally better fit than the linear model. However, for these seven biomarkers, the linear slope estimates extracted from the two models were highly correlated in sex-adjusted models (waist–hip ratio: 0.99, VO_2Max: 1.00, FEV_1/FVC: 0.99, FEV_1: 0.99, apolipoprotein B100/A1 ratio: 0.99; BUN: 0.99; gum health: 0.99), leading us to conclude that we could reasonably use the linear slope estimates from the models including linear fixed effects only. This is graphically apparent in Supplementary Fig. 7, which compares the linear-only and linear + quadratic growth curves.

Structural MRI. *Image acquisition.* Each participant was scanned using a Siemens MAGNETOM Skyra (Siemens Healthcare GmbH) 3 T scanner equipped with a 64-channel head/neck coil at the Pacific Radiology Group imaging center in Dunedin, New Zealand. High-resolution T1-weighted images were obtained using an MP-RAGE sequence with the following parameters: reptition time (TR) = 2,400 ms; echo time (TE) = 1.98 ms; 208 sagittal slices; flip angle, 9°; field of view FOV = 224 mm; matrix = 256 × 256; slice thickness = 0.9 mm with no gap (voxel size 0.9 × 0.875 × 0.875 mm^3) and total scan time = 6 min and 52 s. Three-dimensional fluid-attenuated inversion recovery (FLAIR) images were obtained with the following parameters: TR = 8000 ms; TE = 399 ms; 160 sagittal slices; FOV = 240 mm; matrix = 232 × 256; slice thickness = 1.2 mm (voxel size 0.9 × 0.9 × 1.2 mm^3) and total scan time = 5 min and 38 s. Additionally, a gradient echo field map was acquired with the following parameters: TR = 712 ms; TE = 4.92 and 7.38 ms; 72 axial slices; FOV = 200 mm; matrix = 100 × 100; slice thickness = 2.0 mm (voxel size 2 mm isotropic) and total scan time = 2 min and 25 s. Diffusion-weighted images providing full brain coverage were acquired with 2.5-mm isotropic resolution and 64 diffusion-weighted directions (4,700 ms repetition time, 110.0 ms echo time, *b*-value = 3,000 s mm^{-2}, FOV = 240 mm, 96 × 96 acquisition matrix, slice thickness = 2.5 mm). Nonweighted ($b = 0$) images were acquired in both the encoding (AP) and reverse encoding (PA) directions to allow for echo planar imaging distortion correction. A total of 875 study members completed the MRI scanning protocol (see Supplementary Figs. 1 and 2 for attrition analyses).

Image processing. Structural MRI data were analyzed using the Human Connectome Project (HCP) minimal preprocessing pipeline as detailed elsewhere[62]. Briefly, T1-weighted and FLAIR images were processed through the PreFreeSurfer, FreeSurfer and PostFreeSurfer pipelines. T1-weighted and FLAIR images were corrected for readout distortion using the gradient echo field map, coregistered, brain-extracted and aligned together in the native T1 space using boundary-based registration[63]. Images were then processed with a custom FreeSurfer recon-all pipeline that is optimized for structural MRI with a higher resolution than 1 mm isotropic. Finally, recon-all output was converted into CIFTI (Connectivity Informatics Technology Initiative) format and registered to a common 32k_FS_LR mesh using MSM-sulc[64]. Outputs of the minimal preprocessing pipeline were checked visually for accurate surface generation by examining each participant's myelin map, pial surface and white matter boundaries.

Cortical thickness, surface area and hippocampal volume. For each participant, the mean cortical thickness and surface area were extracted from each of the 360 cortical parcels in the HCP-MPP1.0 parcellation[30]. Average cortical thickness and average surface area (reported in Tables 1 and 2) were calculated as the average value of cortical thickness and surface area across these 360 parcels. Regional cortical thickness and surface area measures have each been found to have excellent test–retest reliability in this sample (mean intraclass correlation coefficients (ICCs) = 0.85 and 0.99 respectively)[65]. Bilateral hippocampal volume was extracted from the FreeSurfer 'aseg' parcellation. Of the 875 study members for whom data were available, 4 were excluded due to major incidental findings or previous injuries (for example, large tumors or extensive damage to the brain/skull), 9 due to missing FLAIR or field map scans and 1 due to poor surface mapping. This resulted in cortical thickness, surface area and hippocampal volume data for 861 study members.

White matter hyperintensities. To identify and extract the total volume of WMH, T1-weighted and FLAIR images for each participant were processed with the UBO Detector, a cluster-based, fully automated pipeline with established out-of-sample performance and high reliability in our data (test–retest ICC = 0.87)[66,67]. The resulting WMH probability maps were thresholded at 0.7, which is the suggested standard. WMH volume is measured in Montreal Neurological Institute space, thus removing the influence of differences in brain volume on WMH volume. Because of the potential for bias and false positives due to the thresholds and masks applied in UBO, the resulting WMH maps for each study member were checked manually by two independent raters to ensure that false detections did not substantially contribute to estimates of WMH volume. Visual inspections were done blind to the participants' cognitive status. Due to the tendency of automated algorithms to mislabel regions surrounding the septum as WMH, these regions were masked manually out to further ensure the most accurate grading possible. WMH data were excluded if study members had missing FLAIR scans, multiple sclerosis, inaccurate white matter labeling or low-quality MRI data, yielding 852 datasets for analyses. In all analyses, WMH volume was log-transformed.

Diffusion-weighted imaging. Diffusion-weighted images were processed in the Oxford Center for Functional MRI of the Brain (FMRIB)'s software library (http://fsl.fmrib.ox.ac.uk/fsl). Raw diffusion-weighted images were corrected for susceptibility artifacts, movement and eddy currents using topup and eddy. Images were then skull-stripped and fitted with diffusion tensor models at each voxel using FMRIB's Diffusion Toolbox (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT). The resulting fractional anisotropy (FA) images from all study members were registered nonlinearly to the FA template developed by the Enhancing Neuro Imaging Genetics Through Meta-Analysis consortium (ENIGMA), a minimal deformation target calculated across a large number of individuals[68]. The images were then processed using the tract-based spatial statistics analytic method[69] modified to project individual FA values onto the ENIGMA–diffusion tensor imaging skeleton. Following the extraction of the skeletonized white matter and projection of individual FA values, ENIGMA-tractwise regions of interest, derived from the Johns Hopkins University white matter parcellation atlas[70], were transferred to extract the mean FA across the full skeleton and average FA values for a total of 25 (partially overlapping) tracts. After visual inspection of all diffusion images, seven study members were removed because data were collected with a 20-channel head coil to accommodate claustrophobia or large head size, leading to poorer diffusion image quality; three were removed due to major incidental findings; five were removed due to excessive (>3 mm) motion detected with eddy tool and seven were removed due to missing diffusion scans. This resulted in diffusion images for 854 study members available for analyses.

Brain age. We estimated brain age with a publicly available algorithm, developed by a different research team, which uses information about cortical anatomy to estimate the age of a person's brain[32]. This algorithm was trained on chronological age in samples ranging from 19 to 82 years old. The algorithm has been shown to predict chronological age in multiple independent samples and to have high test–retest reliability in the Dunedin Study (ICC = 0.81)[71], although it has a documented tendency to underestimate chronological age by approximately 3 years among adults between chronological ages 44 and 46 years[72]. For this reason, we standardized the scores to the mean chronological age of the Dunedin Study members at the time of their scanning in the Phase-45 assessment[72]. In all analyses, we used brainAGE, which is the difference between each Study member's estimated brain age and their chronological age. An older brainAGE results when the predicted brain age is older than the study member's chronological age and is presumed to reflect accelerated brain aging. Data from six study members were excluded due to major incidental findings or previous head injuries (for example, large tumors or extensive damage to the brain or skull). This resulted in brainAGE scores for 869 study members available for analyses.

Cognitive functioning. *Neurocognitive functioning.* The Wechsler Adult Intelligence Scale-IV (WAIS-IV)[73] was administered to each participant at age 45 years, yielding the IQ. In addition to full-scale IQ, the WAIS-IV yields indexes of four specific cognitive function domains: processing speed, working memory, perceptual reasoning and verbal comprehension.

Child-to-adult neurocognitive decline. The Wechsler Intelligence Scale for Children–Revised (WISC–R)[74] was administered to each participant at ages 7, 9 and 11 years, yielding the IQ. To increase baseline reliability, we averaged each participant's three scores. We measured cognitive decline by studying IQ scores at midlife after controlling for IQ scores in childhood (as a sensitivity analysis, in addition to analyzing residualized change, we also analyzed 'change scores' assessed as the difference between adult IQ and childhood IQ and obtained the same substantive and statistically significant results). We focus on change in the overall IQ given evidence that age-related slopes are correlated across all cognitive functions,

indicating that research on cognitive decline may be best focused on a highly reliable summary index, rather than focused on individual functions[75].

RAVL test. The RAVL test is a test of verbal learning and memory administered at 45 years[76]. The test involves repeated presentation of a 15-word list and a one-time presentation of an interference list. Total recall is the total number of words (0–60) recalled over four trials (the sum of words recalled across trials 1–4). Delayed recall is the total number of words (0–15) recalled after a 30-min delay, with interfering cognitive tasks in the interim.

Informant memory and attention. Subjective everyday cognitive function was reported by individuals nominated by each participant as knowing him/her well. These informants were mailed questionnaires and asked to complete a checklist indicating whether the study member had problems with memory or attention over the past year. These questionnaires were designed to be consistent with the 'subjective impression' criteria for mild cognitive impairment from the DSM-V[77]; 94% of study members had at least one informant return the questionnaire, 88% had two and 68% had three. A memory-problems scale consisted of three items: 'has problems with memory,' 'misplaces wallet, keys, eyeglasses and paperwork' and 'forgets to do errands, return calls and pay bills' (internal consistency reliability = 0.63). An attention problems scale consisted of four items: 'is easily distracted, gets sidetracked easily,' 'can't concentrate, mind wanders,' 'tunes out instead of focusing' and 'has difficulty organizing tasks that have many steps' (internal consistency reliability = 0.79). For each question, informants were asked to rate the study member on a 0–2 scale (0 = doesn't apply, 1 = applies somewhat and 2 = definitely applies). Scores were then summed within each rater and averaged across raters.

Sensory–motor functioning. We assessed sensory–motor functional capacity at age 45 with objective tests of physical and sensory functioning and self-reports of physical limitations.

Gait speed. Gait speed ($m\,s^{-1}$) was assessed with the 6-m-long GAITRite Electronic Walkway (CIR Systems, Inc.) with 2-m acceleration and 2-m deceleration before and after the walkway, respectively. Gait speed was assessed under three walking conditions: usual gait speed (walk at a normal pace from a standing start, measured as a mean of two walks) and two challenge paradigms, dual-task gait speed (walk at a normal pace while reciting alternate letters of the alphabet out loud, starting with the letter A, measured as a mean of two walks) and maximum gait speed (walk as fast as safely possible, measured as a mean of three walks). Gait speed was correlated across the three walk conditions[78]. To increase reliability and take advantage of the variation in all three walk conditions (usual gait and the two challenge paradigms), we calculated the mean of the three highly correlated individual walk conditions to generate our primary measure of composite gait speed.

One-legged balance. Balance was measured using the Unipedal Stance Test as the maximum time achieved across three trials of the test with eyes closed[79–81].

Chair-stand test. Chair rises were measured as the number of stands with no hands a participant completed in 30 s from a seated position[82,83].

Two-minute step test. The 2-min step test was measured as the number of times a participant lifted their right knee to midthigh height (measured as the height halfway between the knee cap and the iliac crest) in 2 min at a self-directed pace[83,84].

Grip strength. Handgrip strength was measured for each hand (elbow held at 90°, upper arm held tight against the trunk) as the maximum value achieved across three trials using a Jamar digital dynamometer[37,85].

Visual-motor coordination. Visual-motor coordination was measured as the time to completion of the Grooved Pegboard Test[76]. Scores for the Grooved Pegboard test were reversed so that higher values corresponded to better performance.

Contrast sensitivity. Study members wore their glasses or contact lenses (if these were normally worn). Study members were seated 1 m from the Thomson Test Chart and a Samsung 23 inch LCD Thin Client screen. Room lighting was set at 520 lux. Contrast sensitivity was tested with both eyes open. The Pelli-Robson chart presents three letters per line and the black letters gradually fade from black to gray to white on the white background to determine the lowest level of contrast that the eye can detect. If only one letter on a line was determined correctly by the study member, the number of letters was recorded to determine the contrast sensitivity function score. However, if two letters on a line were determined correctly, the technician proceeded to the next line to determine whether the study member could determine any of these letters correctly.

Audiometry. Hearing acuity was assessed in a sound-attenuating booth (350 Series MaxiAudiology Booth by IAC Acoustics) that met the standard for maximum permissible ambient sound pressure levels. Pure tone audiometry was administered via the Interacoustics Callisto Suite configured to the Interacoustics OtoAccess database, operated from a Hewlett Packard Envy laptop with sound delivered by Sennheiser HDA 300 headphones. The program was set to deliver pure-tone stimuli in the following order: 1,000 Hz, 2,000 Hz, 4,000 Hz, 8,000 Hz, 12,500 Hz and 500 Hz. Presentation intensity levels began at 40 decibels at hearing level (dB HL) for normal-hearing study members and 60 dB HL for hearing-aid users. Audiometry used the Hughson-Westlake procedure (ISO8253–1:2010; Acoustics-Audiometric test methods-Part1: Pure-tone air and bone-conduction audiometry) in which participants respond when they hear a pure tone. Auditory thresholds, defined as the lowest intensity level that the individual responded to, for two out of three presentations, were determined using a standard down-10-up-5 technique for each frequency. A four-frequency pure-tone average was calculated by averaging 500 Hz, 1,000 Hz, 2,000 Hz and 4,000 Hz, and a high pure-tone average was calculated by averaging 8,000 Hz and 12,500 Hz. The results for the 'best ear' are reported.

Spatial listening. Study members completed the LiSN-S (Phonak, Switzerland) in a sound-attenuating booth (350 Series MaxiAudiology Booth by IAC Acoustics). Auditory stimuli were delivered through a pair of Sennheiser 215 headphones attached to a Mini PCM2704 external sound card. The LiSN-S produces a three-dimensional auditory environment through the headphones via four different task conditions[86]. Target sentences are superimposed with distractor stories (maskers). Across the four conditions, these maskers differ with respect to perceived spatial location (0° or ±90° azimuth) and speaker identity (same or different from the target speaker). The following order of conditions was identically presented to all participants: (1) different speaker at ±90° azimuth, (2) same speaker at ±90° azimuth, (3) different speaker at ±0° azimuth and (4) same speaker at ±0° azimuth.

The masking stories were presented consistently at an intensity of 55 decibels sound pressure level (dB SPL). Participants repeated the target sentences and were scored in the software on their accuracy (words correct in each sentence). The program was adaptive, with target sentences delivered at 62 dB SPL to start, and intensity levels continuously adjusted up (if <50% of the words in the sentence correct) and down (if >50% of the words in the sentence correct) on the basis of accuracy. The first few sentences (a minimum of five) are considered practice sentences. This practice testing continues where levels were lowered in 4 dB increments, until one upward reversal in performance was recorded (that is the sentence score drops <50% of words correct), after which the increments decreased to ±2 dB steps. Practice sentence scores did not form part of the final score. The test condition continued until the average of the levels from positive-and negative-going reversals amounted to ≥3 (independent midpoint target level), and the standard error of these midpoints was less than 1 dB. Alternatively, the test condition continued until it reached the maximum number of 30 sentence presentations. Speech reception thresholds were calculated as the lowest intensity at which the individual could repeat 50% of the words correctly. Two outcome scores were used: (1) speech reception threshold from a low-cue condition represented performance in the most difficult auditory environment (masker speaker same as the target speaker, and masker was presented at 0° azimuth, in the same location as the target speaker). (2) Spatial advantage score measured the benefit gained when the masker is presented from a different direction from the target.

Physical limitations. Physical limitations were measured with the RAND 36-Item Health Survey v.1.0 physical functioning scale[87]. Participant responses ('limited a lot', 'limited a little' or 'not limited at all') assessed their difficulty with completing various activities, for example, climbing several flights of stairs, walking more than 1 km, participating in strenuous sports, etc. Scores were reversed to reflect physical limitations so that a high score indicates more limitations.

Perceptions of aging. *Self, informant and researcher impressions.* Across several measures (described below), perceptions of aging were assessed through self, informant and researcher ratings. To strengthen our insight into the subjective perceptions of aging we adopted a multi-informant approach by triangulating amongst several reporters. Informants tended to be friends and family who knew the study members well. Therefore informants often had access to at least some information about study members' health histories. Staff ratings were obtained from four raters for each study member: the cardiovascular nurses, the sensory technicians, the study director and the assessment manager (who was in charge of informed consent and logistics on the in-Unit assessment day). During the assessment day, the cardiovascular nurses became aware of the participants' blood pressure and their performance on the exercise bike during their session, but had no knowledge of other aspects of the study members' health status or health history. The sensory technicians became aware of the participants' performance on vision and hearing tests but had no other information about the study members' health status or health history. The director and the assessment manager did not have access to information about current health on the assessment day or the study member's health history.

Attitudes towards aging. Age beliefs were assessed with the five-item Attitude Toward Aging scale[41]. Sample items include 'Things keep getting worse as I get older (R)' and 'As you get older, you are less useful.'

Perceived health. We obtained three reports about study members' health from three sources: self-reports, informant impressions and staff impressions (see next paragraph for a description of these data sources). All reporters rated the study member's general health using the following response options: excellent, very good, good, fair or poor. Correlations between self-, informant and staff ratings ranged from 0.48 to 0.55.

Age appearance. We obtained reports about study members' age appearance from three sources: self-reports, informant impressions and staff impressions. For self-reports, we asked study members about their own impressions of how old they looked: 'Do you think you LOOK older, younger or about your actual age?' Response options were younger than their age, about their actual age or older than their age. We also asked study members to rate their age perceptions in years: 'How old do you feel?'. For informant impressions, informants who knew a study member well (94% response rate) were asked: 'Compared to others their age, do you think he/she (the study member) looks younger or older than others their age?'. Response options were: much younger, a bit younger, about the same, a bit older or much older. To collect staff impressions, four members of the Dunedin Study Unit staff completed a brief questionnaire describing each study member. To assess age appearance, staff used a 7-item scale to assign a 'relative age' to each study member (1 = young looking, 7 = old looking). Correlations between self-, informant and staff ratings ranged from 0.34 to 0.52.

Facial age. Facial age was on the basis of ratings by an independent panel of eight raters of each participant's digital facial photograph. Facial age was on the basis of two measurements of perceived age. First, age range was assessed by an independent panel of four raters, who were presented with standardized (nonsmiling) facial photographs of participants and were kept blind to their actual age. Raters used a Likert scale to categorize each participant into a 5-year age range (that is, from 20–24 years old up to 70+ years old) (interrater reliability = 0.77). Scores for each participant were averaged across all raters. Second, relative age was assessed by a different panel of four raters, who were told that all photos were of people aged 45 years old. Raters then used a 7-item Likert scale to assign a 'relative age' to each participant (1 = young looking, 7 = old looking) (interrater reliability = 0.79). The measure of perceived age at 45 years, facial age, was derived by standardizing and averaging age range and relative age scores. Visual representations of facial age (Fig. 4b) were created by averaging together facial images of Dunedin Study members. Each of the four images in Fig. 4b was created by averaging together ten facial photos using the Psychomorph software (https://webmorph.org)[88]. Ten study members' facial images went into each average male and female face. The four groups of ten images were chosen on the basis of their facial age ratings (youngest ten females, youngest ten males, oldest ten females and oldest ten males).

Perceived longevity. At age 45, study members were asked, 'How likely is it that you will live to be 75 or more?' (0 = not likely, 1 = somewhat likely or 2 = very likely).

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

Data availability

The Dunedin Study datasets reported in the current Article are not publicly available due to a lack of informed consent and ethical approval for public data sharing. The Dunedin study datasets are available on request by qualified scientists. Requests require a concept paper describing the purpose of data access, ethical approval at the applicant's university and provision for secure data access (https://moffittcaspi.trinity.duke.edu/research-topics/dunedin). We offer secure access on the Duke, Otago and King's College campuses.

Code availability

Custom code that supports the findings of this study is available from the corresponding author on request.

Received: 27 July 2020; Accepted: 10 February 2021;
Published online: 15 March 2021

References

1. Vos, T. et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2163–2196 (2012).
2. Crimmins, E. M. Lifespan and healthspan: past, present, and promise. *Gerontologist* **55**, 901–911 (2015).
3. Kaeberlein, M. Longevity and aging. *F1000Prime Rep.* **5**, 5 (2013).
4. Kirkwood, T. B. L. Understanding the odd science of aging. *Cell* **120**, 437–447 (2005).
5. Gladyshev, V. N. Aging: progressive decline in fitness due to the rising deleteriome adjusted by genetic, environmental, and stochastic processes. *Aging Cell* **15**, 594–602 (2016).
6. Kennedy, B. K. et al. Geroscience: linking aging to chronic disease. *Cell* **159**, 709–713 (2014).
7. Kaeberlein, M., Rabinovitch, P. S. & Martin, G. M. Healthy aging: the ultimate preventative medicine. *Science* **350**, 1191–1193 (2015).
8. Ferrucci, L. & Fabbri, E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. *Nat. Rev. Cardiol.* **15**, 505–522 (2018).
9. Melzer, D., Pilling, L. C. & Ferrucci, L. The genetics of human ageing. *Nat. Rev. Genet.* **21**, 88–101 (2020).
10. Chen, B. H. et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging* **8**, 1844–1865 (2016).
11. Parker, D. C. et al. Association of blood chemistry quantifications of biological aging with disability and mortality in older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* **16**, 1671–1679 (2020).
12. Levine, M. E. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *J. Gerontol. A Biol. Sci. Med. Sci.* **68**, 667–674 (2013).
13. Belsky, D. W. et al. Eleven telomere, epigenetic clock, and biomarker-composite quantifications of biological aging: do they measure the same thing? *Am. J. Epidemiol.* **187**, 1220–1230 (2018).
14. Li, Q. et al. Homeostatic dysregulation proceeds in parallel in multiple physiological systems. *Aging Cell* **14**, 1103–1112 (2015).
15. Sebastiani, P. et al. Biomarker signatures of aging. *Aging Cell* **16**, 329–338 (2017).
16. Levine, M. E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
17. Sierra, F. Geroscience and the challenges of aging societies. *Aging Med.* **2**, 132–134 (2019).
18. Fahy, G. M. et al. Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell* **18**, e13028 (2019).
19. Belsky, D. W., Huffman, K. M., Pieper, C. F., Shalev, I. & Kraus, W. E. Change in the rate of biological aging in response to caloric restriction: CALERIE Biobank. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **73**, 4–10 (2017).
20. Moffitt, T. E., Belsky, D. W., Danese, A., Poulton, R. & Caspi, A. The longitudinal study of aging in human young adults: knowledge gaps and research agenda. *J. Gerontol. A. Biol. Sci. Med. Sci.* **72**, 210–215 (2017).
21. Poulton, R., Moffitt, T. E. & Silva, P. A. The Dunedin multidisciplinary health and development study: overview of the first 40 years, with an eye to the future. *Soc. Psychiatry Psychiatr. Epidemiol.* **50**, 679–693 (2015).
22. Belsky, D. W. et al. Quantification of biological aging in young adults. *Proc. Natl Acad. Sci. USA* **112**, E4104–E4110 (2015).
23. Wilson, R. S., Leurgans, S. E., Boyle, P. A., Schneider, J. A. & Bennett, D. A. Neurodegenerative basis of age-related cognitive decline. *Neurology* **75**, 1070–1078 (2010).
24. Feigin, V. L. et al. Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol.* **18**, 459–480 (2019).
25. Elliott, M. L. MRI-Based biomarkers of accelerated aging and dementia risk in midlife: how close are we?. *Ageing Res. Rev.* **61**, 101075 (2020).
26. Sperling, R., Mormino, E. & Johnson, K. The evolution of preclinical Alzheimer's disease: implications for prevention trials. *Neuron* **84**, 608–622 (2014).
27. Rensma, S. P., van Sloten, T. T., Launer, L. J. & Stehouwer, C. D. A. Cerebral small vessel disease and risk of incident stroke, dementia and depression, and all-cause mortality: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* **90**, 164–173 (2018).
28. Franke, K. & Gaser, C. Longitudinal changes in individual BrainAGE in healthy aging, mild cognitive impairment, and Alzheimer's disease. *GeroPsych (Bern)* **25**, 235–245 (2012).
29. Fjell, A. M. et al. Accelerating cortical thinning: unique to dementia or universal in aging? *Cereb. Cortex* **24**, 919–934 (2014).
30. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
31. Nyberg, L. & Pudas, S. Successful memory aging. *Annu. Rev. Psychol.* **70**, 219–243 (2019).
32. Liem, F. et al. Predicting brain-age from multimodal imaging data captures cognitive impairment. *Neuroimage* **148**, 179–188 (2017).
33. Tucker-Drob, E. M. Cognitive aging and dementia: a life-span. *Perspect. Annu. Rev. Dev. Psychol.* **1**, 177–196 (2019).
34. Whalley, L. J. et al. Childhood mental ability and dementia. *Neurology* **55**, 1455–1459 (2000).
35. Knudtson, M. D., Klein, B. E. K. & Klein, R. Biomarkers of aging and falling: the beaver dam eye study. *Arch. Gerontol. Geriatr.* **49**, 22–26 (2009).
36. Studenski, S. et al. Gait speed and survival in older adults. *J. Am. Med. Assoc.* **305**, 50–58 (2011).
37. Rantanen, T. et al. Midlife hand grip strength as a predictor of old age disability. *J. Am. Med. Assoc.* **281**, 558–560 (1999).
38. Genther, D. J. et al. Association of hearing impairment and mortality in older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* **70**, 85–90 (2015).
39. Lin, F. R. et al. Hearing loss and cognition in the Baltimore longitudinal study of aging. *Neuropsychology* **25**, 763–770 (2011).

40. Rippon, I. & Steptoe, A. Feeling old vs being old: associations between self-perceived age and mortality. *JAMA Intern. Med.* **175**, 307–309 (2015).
41. Levy, B. R., Slade, M. D. & Kasl, S. V. Longitudinal benefit of positive self-perceptions of aging on functional health. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **57**, 409–417 (2002).
42. Brown, R. T. & Covinsky, K. E. Moving prevention of functional impairment upstream: is middle age an ideal time for intervention? *Women's Midlife Heal.* **6**, https://doi.org/10.1186/s40695-020-00054-z (2020).
43. Brunner, E. J. et al. Midlife contributors to socioeconomic differences in frailty during later life: a prospective cohort study. *Lancet Public Heal.* **3**, E313–E222 (2018).
44. Livingston, G. et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* **396**, 413–446 (2020).
45. Barzilai, N., Crandall, J. P., Kritchevsky, S. B. & Espeland, M. A. Metformin as a tool to target aging. *Cell Metab.* **23**, 1060–1065 (2016).
46. Justice, J. N. & Kritchevsky, S. B. Putting epigenetic biomarkers to the test for clinical trials. *eLife* **9**, e54870 (2020).
47. Schaie, K. W. Age changes and age differences. *Gerontologist* **7**, 128–132 (1967).
48. Belsky, D. W., Ma, J., Cohen, A. A., Griffith, L. E. & Raina, P. Comparing biological age estimates using domain-specific measures from the Canadian longitudinal study on aging. *J. Gerontol. A Biol. Sci. Med. Sci.* **76**, 187–194 (2021).
49. Feng, D., Silverstein, M., Giarrusso, R., McArdle, J. J. & Bengtson, V. L. Attrition of older adults in longitudinal surveys: detection and correction of sample selection bias using multigenerational data. *J. Gerontol. B Psychol Sci. Soc. Sci.* **61**, S323–S328 (2006).
50. Goudy, W. J. Sample attrition and multivariate analysis in the retirement history study. *J. Gerontol.* **40**, 358–367 (1985).
51. Belsky, D. W. et al. Quantification of the pace of biological aging in humans through a blood test, the DunedinPoAm DNA methylation algorithm. *eLife* **9**, e54870 (2020).
52. Funder, D. C. & Ozer, D. J. Evaluating effect size in psychological research: sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* **2**, 156–168 (2019).
53. Hofer, S. M., Sliwinski, M. J. & Flaherty, B. P. Understanding ageing: further commentary on the limitations of cross-sectional designs for ageing research. *Gerontology* **48**, 22–29 (2002).
54. Goldman, D. P. et al. Substantial health and economic returns from delayed aging may warrant a new focus for medical research. *Health Aff.* **32**, 1698–1705 (2013).
55. Moffitt, T. E. Behavioral and social research to accelerate the geroscience translation agenda. *Ageing Res. Rev.* **63**, 101146 (2020).
56. Sierra, F. et al. The role of multidisciplinary team science in overcoming barriers to moving geroscience from the bench to clinical care and health policy. *J. Am. Geriatr. Soc.* (in the press).
57. Arias, E. United States Life Tables, 2017. *National Vital Statistics Reports* **68**, https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_07-508.pdf (2019).
58. Song, Z. Potential implications of lowering the medicare eligibility age to 60. *J. Am. Med. Assoc.* **323**, 2472–2473 (2020).
59. Moffitt, R. A. & Ziliak, J. P. Entitlements: options for reforming the social safety net in the United States. *Ann. Am. Acad. Polit. Soc. Sci.* **686**, 8–35 (2019).
60. Richmond-Rakerd, L. S. et al. Clustering of health, crime and social-welfare inequality in 4 million citizens from two nations. *Nat. Hum. Behav.* **44**, 255–264 (2020).
61. Thomas, S. Telomeres as sentinels for environmental exposures, psychosocial stress, and disease susceptibility. In *Workshop Summary. A Workshop Co-sponsored by the National Institute of Environmental Health Sciences (NIEHS) and the National Institute on Aging (NIA)* (Rose Li and Associates, 2017).
62. Glasser, M. F. et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage* **80**, 105–124 (2013).
63. Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* **48**, 63–72 (2009).
64. Robinson, E. C. et al. MSM: a new flexible framework for multimodal surface matching. *Neuroimage* **100**, 414–426 (2014).
65. Elliott, M. L. et al. What is the test–retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychol. Sci.* **31**, 792–806 (2020).
66. Jiang, J. et al. UBO detector – a cluster-based, fully automated pipeline for extracting white matter hyperintensities. *Neuroimage* **174**, 539–549 (2018).
67. d'Arbeloff, T. et al. White matter hyperintensities are common in midlife and already associated with cognitive decline. *Brain Commun.* **1**, fcz041 (2019).
68. Jahanshad, N. et al. Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: a pilot project of the ENIGMA-DTI working group. *Neuroimage* **81**, 455–469 (2013).
69. Smith, S. M. et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage* **31**, 1487–1505 (2006).
70. Mori, S. et al. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage* **40**, 570–582 (2008).
71. Elliott, M. L. et al. Brain-age in midlife is associated with accelerated biological aging and cognitive decline in a longitudinal birth cohort. *Mol. Psychiatry* https://doi.org/10.1038/s41380-019-0626-7 (2019).
72. Liang, H., Zhang, F. & Niu, X. Investigating systematic bias in brain age estimation with application to post-traumatic stress disorders. *Hum. Brain Mapp.* **40**, 3143–3152 https://doi.org/10.1002/hbm.24588 (2019).
73. Wechsler, D. *Wechsler Adult Intelligence Scale* 4th edn (Pearson Assessment, 2008).
74. Wechsler, D. *Manual for the Wechsler Intelligence Scale for Children, Revised* (Psychological Corporation, 1974).
75. Tucker-Drob, E. M., Brandmaier, A. M. & Lindenberger, U. Coupled cognitive changes in adulthood: a meta-analysis. *Psychol. Bull.* **145**, 273–301 (2019).
76. Lezak, D, Howieson, D, Loring, D, Hannay, H. & Fischer, J. *Neuropsychological Assessment* 4th edn (Oxford Univ. Press, 2004).
77. Ganguli, M. Can the DSM-5 framework enhance the diagnosis of MCI? *Neurology* **81**, 2045–2050 (2013).
78. Rasmussen, L. J. H. et al. Association of neurocognitive and physical function with gait speed in midlife. *JAMA Netw. Open* **2**, e1913123 (2019).
79. Bohannon, R. W., Larkin, P. A., Cook, A. C., Gear, J. & Singer, J. Decrease in timed balance test scores with aging. *Phys. Ther.* **64**, 1067–1070 (1984).
80. Vereeck, L., Wuyts, F., Truijen, S. & Van de Heyning, P. Clinical assessment of balance: normative data, and gender and age effects. *Int. J. Audiol.* **47**, 67–75 (2008).
81. Springer, B. A., Marin, R., Cyhan, T., Roberts, H. & Gill, N. W. Normative values for the unipedal stance test with eyes open and closed. *J. Geriatr. Phys. Ther.* **30**, 8–15 (2007).
82. Jones, C. J., Rikli, R. E. & Beam, W. C. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Res. Q. Exerc. Sport* **70**, 113–119 (1999).
83. Jones, C. J. & Rikli, R. E. Measuring functional fitness of older adults. *J. Act. Aging* **2002**, 24–30 (2002).
84. Rikli, R. E. & Jones, C. J. Functional fitness normative scores for community-residing older adults, ages 60–94. *J. Aging Phys. Act.* **7**, 162–181 (1999).
85. Mathiowetz, V. et al. Grip and pinch strength: normative data for adults. *Arch. Phys. Med. Rehabil.* **66**, 69–74 (1985).
86. Cameron, S. & Dillon, H. Development of the listening in spatialized noise-sentences test (LISN-S). *Ear Hear.* **28**, 196–211 (2007).
87. RAND 36-Item Short Form Survey (SF-36); https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html
88. DeBruine, L. debruine/webmorph: Beta release v.2 (Zenodo, 2018); https://doi.org/10.5281/ZENODO.1162670

Acknowledgements

This research was supported by the National Institute on Aging (NIA) grant nos. R01AG032282 and R01AG049789, and the UK Medical Research Council grant no. MR/P005918/1. Additional support was provided by the Jacobs Foundation, grant nos. NIA P30 AG028716 and NIA P30 AG034424. The Dunedin Multidisciplinary Health and Development Research Unit was supported by the New Zealand Health Research Council (Project Grant nos. 15-265 and 16-604) and the New Zealand Ministry of Business, Innovation and Employment (MBIE). M.L.E. is supported by the National Science Foundation Graduate Research Fellowship (no. NSF DGE-1644868) and grant no. NIA F99 AG068432-01. We thank members of the Advisory Board for the Dunedin Neuroimaging Study, Dunedin Study members, Unit research staff, Pacific Radiology Group staff and study founder P. Silva.

Author contributions

M.L.E., A.C., A.R.H., R.P. and T.E.M. designed the research. M.L.E., A.C., R.M.H, A.A., J.M.B., R.J.H., H.L.H., S.H., R.K., A.K., J.H.L., T.R.M., S.C.P., S.R., L.S.R.-R., A.R., K.S., W.M.T., P.R.T., B.S.W., G.W., A.R.H., R.P. and T.E.M. performed the research. M.L.E., R.M.H. and A.K. analyzed data. M.L.E., A.C., A.R.H. and T.E.M. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43587-021-00044-4.

Correspondence and requests for materials should be addressed to M.L.E.

Peer review information *Nature Aging* thanks William Jagust, Juulia Jylhava, Diana Kuh, John Rowe and Thomas Travison for their contribution to the peer review of this work.

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

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

nature research

Corresponding author(s): Maxwell Elliott

Last updated by author(s): Feb 3, 2021

Reporting Summary

Nature Research 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 Research 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 *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	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 *Give P values as exact values whenever suitable.*
☒	☐	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	No software was used.
Data analysis	All statistical analyses were completed using linear regression models in R (version 3.4.0).

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 Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Data are available via managed data access (https://moffittcaspi.trinity.duke.edu/research).

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

Behavioural & social sciences study design

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

Study description	Participants are members of the Dunedin Study, a longitudinal investigation of health and behavior in a representative birth cohort
Research sample	The 1037 participants (91% of eligible births) were all individuals born between April 1972 and March 1973 in Dunedin, New Zealand, who were eligible on the basis of residence in the province and who participated in the first assessment at age 3 years.
Sampling strategy	All individuals born between April 1972 and March 1973 in Dunedin, New Zealand, who were eligible on the basis of residence in the province and who participated in the first assessment at age 3 years.
Data collection	Each participant was scanned using a Siemens MAGNETOM Skyra (Siemens Healthcare GmbH) 3T scanner equipped with a 64-channel head/neck coil at the Pacific Radiology Group imaging center in Dunedin, New Zealand. Neurocognitive functioning. The Wechsler Adult Intelligence Scale-IV (WAIS-IV)67 was administered to each participant at age 45 years, yielding the IQ. Child-to-adult neurocognitive decline. The Wechsler Intelligence Scale for Children–Revised (WISC–R)68 was administered to each participants at ages 7, 9, and 11 years, yielding the IQ. Rey Auditory Verbal Learning Test. This is a test of verbal learning and memory administered at 45 years. Informant memory and attention. Subjective everyday cognitive function was reported by individuals nominated by each participant as knowing him/her well. These informants were mailed questionnaires and asked to complete a checklist indicating whether the Study member had problems with memory or attention over the past year. Gait speed. Gait speed (meters per second) was assessed with the 6-m-long GAITRite Electronic Walkway (CIR Systems, Inc) with 2-m acceleration and 2-m deceleration before and after the walkway, respectively. One-legged balance. Balance was measured using the Unipedal Stance Test as the maximum time achieved across three trials of the test with eyes closed. Chair-stand test. Chair rises were measured as the number of stands with no hands a participant completed in 30 seconds from a seated position 2 min step test. The 2-min step test was measured as the number of times a participant lifted their right knee to mid-thigh height (measured as the height half-way between the knee cap and the iliac crest) in 2 minutes at a self-directed pace. Grip strength. Handgrip strength was measured for each hand (elbow held at 90°, upper arm held tight against the trunk) as the maximum value achieved across three trials using a Jamar digital dynamometer. Visual-motor coordination. Visual-motor coordination was measured as the time to completion of the Grooved Pegboard Test. Contrast Sensitivity. Study members wore their glasses or contact lenses (if these were normally worn). Study members were seated one meter from the Thomson Test Chart and the Samsung 23" LCD Thin Client screen. Audiometry. Hearing acuity was assessed in a sound-attenuating booth (350 Series MaxiAudiology Booth by IAC Acoustics) which met the standard for maximum permissible ambient sound pressure levels. Pure tone audiometry was administered via the Interacoustics Callisto Suite configured to the Interacoustics OtoAccess database, operated from an HP Envy laptop with sound delivered by Sennheiser HDA 300 headphones. Spatial listening. Study members completed the Listening in Spatialised Noise–Sentences Test (LiSN-S) (Phonak, Switzerland) in a sound-attenuating booth (350 Series MaxiAudiology Booth by IAC Acoustics). Auditory stimuli were delivered through a pair of Sennheiser 215 headphones attached to a Mini PCM2704 external sound card. Physical limitations. Physical limitations were measured with the RAND 36-Item Health Survey 1.0 physical functioning scale.
Timing	Assessments were performed at birth; at ages 3, 5, 7, 9, 11, 13, 15, 18, 21, 26, 32, and 38 years; and, most recently (completed April 2019), at age 45 years, when 938 of the 997 participants (94.1%) still alive participated.

Data exclusions	Data were excluded based on quality control. This decision was made separately for each measure and this is documented in the main text.
Non-participation	At age 45 years, when 938 of the 997 participants (94.1%) still alive participated.
Randomization	There was no randomization.

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
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	See above.
Recruitment	See above.
Ethics oversight	Written informed consent was obtained from cohort participants, and study protocols were approved by the institutional ethical review boards of the participating universities.

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

Magnetic resonance imaging

Experimental design

Design type	Structural MRI.
Design specifications	Structural MRI.
Behavioral performance measures	Structural MRI.

Acquisition

Imaging type(s)	Structural
Field strength	3
Sequence & imaging parameters	High resolution T1-weighted images were obtained using an MP-RAGE sequence with the following parameters: TR = 2400 ms; TE = 1.98 ms; 208 sagittal slices; flip angle, 9°; FOV, 224 mm; matrix =256×256; slice thickness = 0.9 mm with no gap (voxel size 0.9×0.875×0.875 mm); and total scan time = 6 min and 52 s. 3D fluid-attenuated inversion recovery (FLAIR) images were obtained with the following parameters: TR = 8000 ms; TE = 399 ms; 160 sagittal slices; FOV = 240 mm; matrix = 232×256; slice thickness = 1.2 mm (voxel size 0.9×0.9×1.2 mm); and total scan time = 5 min and 38 s. Additionally, a gradient echo field map was acquired with the following parameters: TR = 712 ms; TE = 4.92 and 7.38 ms; 72 axial slices; FOV = 200 mm; matrix = 100×100; slice thickness = 2.0 mm (voxel size 2 mm isotropic); and total scan time = 2 min and 25 s
Area of acquisition	Whole brain scan.

Diffusion MRI ☒ Used ☐ Not used

Parameters	Diffusion-weighted images providing full brain coverage were acquired with 2.5 mm isotropic resolution and 64 diffusion weighted directions (4700 ms repetition time, 110.0 ms echo time, b value 3,000 s/mm2, 240 mm field of view, 96×96 acquisition matrix, slice

thickness=2.5 mm). Non-weighted (b=0) images were acquired in both the encoding (AP) and reverse encoding (PA) directions to allow for EPI distortion correction.

Preprocessing

Preprocessing software: Diffusion weighted images were processed in FSL (http://fsl.fmrib.ox.ac.uk/fsl)

Normalization: fractional anisotropy (FA) images from all Study members were non-linearly registered to the FA template developed by the Enhancing Neuro Imaging Genetics Through Meta-Analysis consortium (ENIGMA), a minimal deformation target calculated across a large number of individuals

Normalization template: FA template developed by the Enhancing Neuro Imaging Genetics Through Meta-Analysis consortium (ENIGMA)

Noise and artifact removal: Raw diffusion weighted images were corrected for susceptibility artifacts, subject movement, and eddy currents using topup and eddy. Images were then skull-stripped and fitted with diffusion tensor models at each voxel using FMRIB's Diffusion Toolbox (FDT; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT)

Volume censoring: 5 Study members were removed due to excessive (>3mm) motion detected with eddy tool.

Statistical modeling & inference

Model type and settings: Linear Regression

Effect(s) tested: Linear regression for effects of the Pace of Aging.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s): Hippocampus defined with Freesurfer ASEG

Statistic type for inference (See Eklund et al. 2016): FDR-correction across each parcel.

Correction: FDR

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis