Patterns of multi-domain cognitive aging in participants of the Long Life Family Study

Paola Sebastiani · Stacy L. Andersen · Benjamin Sweigart · Mengtian Du ·
Stephanie Cosentino · Bharat Thyagarajan · Kaare Christensen · Nicole Schupf ·
Thomas T Perls

Received: 16 April 2020 / Accepted: 8 May 2020 / Published online: 8 June 2020
© American Aging Association 2020

Abstract Maintaining good cognitive function at older age is important, but our knowledge of patterns and predictors of cognitive aging is still limited. We used Bayesian model-based clustering to group 5064

participants of the Long Life Family Study (ages 49–110 years) into clusters characterized by distinct trajectories of cognitive change in the domains of episodic memory, attention, processing speed, and verbal

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11357-020-00202-3>) contains supplementary material, which is available to authorized users.

P. Sebastiani (✉) · B. Sweigart · M. Du
Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Avenue, Boston, MA 02118, USA
e-mail: sebas@bu.edu

B. Sweigart
e-mail: sweigart@bu.edu

M. Du
e-mail: mandydu@bu.edu

S. L. Andersen · T. T. Perls
Department of Medicine, Geriatrics Section, Boston University School of Medicine, Robinson 2400, 72 E Concord St, Boston, MA 02118, USA

S. L. Andersen
e-mail: stacy@bu.edu

T. T. Perls
e-mail: thperls@bu.edu

S. Cosentino · N. Schupf
Department of Neurology, Taub Institute for Research on Alzheimer's Disease and the Aging Brain, and the Gertrude H. Sergievsky Center, Columbia University Medical Center, 630

West 168th Street, New York, NY 10032, USA

S. Cosentino
e-mail: sc2460@cumc.columbia.edu

N. Schupf
e-mail: ns24@cumc.columbia.edu

B. Thyagarajan
Department of Laboratory Medicine and Pathology, University of Minnesota Medical School, MMC 609 Mayo, 420 Delaware, Minneapolis, MN 55455, USA
e-mail: thya0003@umn.edu

K. Christensen
Department of Public Health, The Danish Aging Research Center and The Danish Twin Registry, Epidemiology, Biostatistics and Biodemography, University of Southern Denmark, 5000 Odense, Denmark
e-mail: KChristensen@health.sdu.dk

N. Schupf
Department of Epidemiology, Sergievsky Center, Columbia University Mailman School of Public Health, 630 West 168th Street, New York, NY 10032, USA

fluency. For each domain, we identified 4 or 5 large clusters with representative patterns of change ranging from rapid decline to exceptionally slow change. We annotated the clusters by their correlation with genetic and molecular biomarkers, non-genetic risk factors, medical history, and other markers of aging to discover correlates of cognitive changes and neuroprotection. The annotation analysis discovered both predictors of multi-domain cognitive change such as gait speed and predictors of domain-specific cognitive change such as IL6 and NTproBNP that correlate only with change of processing speed or *APOE* genotypes that correlate only with change of processing speed and logical memory. These patterns also suggest that cognitive decline starts at young age and that maintaining good physical function correlates with slower cognitive decline. To better understand the agreement of cognitive changes across multiple domains, we summarized the results of the cluster analysis into a score of cognitive function change. This score showed that extreme patterns of change affecting multiple cognitive domains simultaneously are rare in this study and that specific signatures of biomarkers of inflammation and metabolic disease predict severity of cognitive changes. The substantial heterogeneity of change patterns within and between cognitive domains and the net of correlations between patterns of cognitive aging and other aging traits emphasizes the importance of measuring a wide range of cognitive functions and the need for studying cognitive aging in concert with other aging traits.

Keywords Cognition · Neuropsychology · Aging · Biomarker

Introduction

The association between aging and changes in cognitive function is well documented in the scientific literature and has been mapped to decreases in brain volume that begin in the early 20s, shrinking of the brain cortex in the early 40s, and modifications of neurons and their interconnectivity (Harada et al. 2013; Murman 2015). The rate of decline in cognitive domains accelerates in most older individuals in their 80s and 90s compared with adults in their 60s and 70s (Small et al. 2011; Wilson et al. 2002), with the exception of some domains such as attentional capacity and vocabulary that show

evidence of stability or even gains in function with older age (Goh et al. 2012; Small et al. 2011).

Although older age is generally associated with worsening memory, reasoning, and processing speed (Salthouse 2019), there is significant variability in the onset and rate of decline of these cognitive functions across older adults of the same age: some individuals experience rapid declines, some experience gradual declines, and some remain stable or they slightly improve (Wilson et al. 2002). The patterns of change appear to also differ across cognitive domains in the same individual, and, for example, decline of executive functions may not be associated with changes of memory (Lovden et al. 2005). The heterogeneity of the patterns of change within and between cognitive domains appears to be important since it has been related to the likelihood of progression to Alzheimer's disease (Cloutier et al. 2015). In addition, while most of the literature on cognitive aging focuses on detecting trajectories of normal cognitive decline (Salthouse 2019), and diverging patterns that can be predictive of disease, dementia and Alzheimer's (Harada et al. 2013), studies of centenarians have shown that maintaining good cognitive function at very old age is also possible (Andersen et al. 2012), and it would be of interest to describe such patterns of exceptionally healthy cognitive aging.

Discovering the most common patterns of change of specific cognitive domains, characterizing how they correlate with each other and with other physical and biological markers of aging, and whether they are predictive of morbidity and mortality could improve our understanding of the effect of aging on cognitive function changes and point to more accurate and comprehensive ways of measuring “cognitive age.” However, there are two challenges to this goal. One challenge is to identify data sets with comprehensive longitudinal assessments of cognitive function that include a variety of patterns of decline and also a wide range of molecular and environmental risk factors. The second challenge is to choose the appropriate statistical analysis to discover patterns of change rather than summary trajectories. In this work, we resolved these two challenges by implementing a Bayesian machine learning analysis of neuropsychological tests that were administered to participants of the Long Life Family Study (LLFS).

Between 2006 and 2009, the Long Life Family Study (LLFS) enrolled approximately 5000 individuals from longevous families and their spouses with the goal of discovering factors that predispose to

healthy aging (Newman et al. 2011). Over the years, this study has collected longitudinal data on cognitive function through many neuropsychological tests, in addition to genetic and molecular biomarkers, clinical events, and markers of physical function. LLFS data have shown that members of longevous families outperform their spouses on many aging traits including, in particular, prevalence of cognitive impairment and dementia (Barral et al. 2013; Cosentino et al. 2013; Sebastiani et al. 2013). However, in-depth analyses have also shown that there is substantial heterogeneity in their healthy aging phenotypes (Marron et al. 2019), thus suggesting that this cohort may provide access to a wide variety of patterns of cognitive decline in multiple domains as well as patterns of exceptionally healthy cognitive aging. To discover these patterns, we used a machine learning technique known as Bayesian model-based clustering to analyze the scores of neuropsychological tests that were administered to assess episodic memory, attention, processing speed, and verbal fluency. This machine learning technique is tailored to discover clusters of participants who share patterns of change that differ from other clusters of participants (Lunn 2013). While a mixed-effect regression—the standard analysis of this type of data—would identify one single trajectory per test that represents the “average” pattern of decline (Salthouse 2019), model-based clustering discovers multiple patterns of decline that are statistically different. The use of this approach to discover patterns of cognitive decline is growing, and it is proving to be useful for stratification of risk of cognitive impairment and Alzheimer’s (Blanken et al. 2020; Lee et al. 2018). To assess the clinical and biological value of the patterns of change discovered with this analysis, we annotated the patterns by their correlation with patients’ medical history, medications, circulating biomarkers, genetic markers, and changes of other aging traits and clinical status over time. The analysis shows that there is substantial heterogeneity in how different people decline in the same cognitive domain but also in how the same person declines across different cognitive domains. The heterogeneity of the decline of cognitive function makes its prediction harder, but our analysis shows that there are genetic and non-genetic biomarkers that can inform about decline in specific cognitive domains.

Materials and methods

Study population

The Long Life Family Study (LLFS) is a family-based study of healthy aging and longevity that enrolled 583 families selected for familial longevity and approximately 5000 family members during 2006 to 2009 (Newman et al. 2011; Sebastiani et al. 2009) and an additional 151 between 2014 and 2017. Participants were enrolled via three American field centers (Boston, Pittsburgh, and New York) and a European field center in Denmark. Potential probands (=index cases) were identified through Medicare enrollees and selected for further screening based on older age and capacity to understand the study (Newman et al. 2011). The eligibility of their family to enter the study was based on the Family Longevity Selection Score (FLoSS) that quantifies the degree of familial longevity using sex and birth-year cohort survival probabilities of the proband and their siblings (Sebastiani et al. 2009). Eligibility of the family to the study was based on a FLoSS score > 7 and having at least one living sibling of the proband and at least one offspring willing to be enrolled in the study. Socio-demographic, medical history data, current medical conditions and medications, physical and cognitive function data, and blood samples were collected via in-person visits and phone questionnaires for all subjects at the time of enrollment as described elsewhere (Newman et al. 2011), and the age of the oldest participants was validated (Elo et al. 2013). Patients are followed up annually to track vital and health status and approximately 80% of surviving participants consented to a second in-person evaluation between 2014 and 2017. Genome-wide genotype data were generated using Illumina SNP array (Bae et al. 2013), and genome-wide genotype data are available from dbGaP (dbGaP Study Accession: phs000397.v1.p1). Approximately 40 circulating biomarkers were measured in all participants as described in (Sebastiani et al. 2016). All subjects provided informed consent approved by the field centers’ institutional review board (IRB).

Neuropsychological tests

At the baseline assessment, participants completed eight of the ten tests in the National Alzheimer’s Coordinating Center (NACC) Unified Data Set (Weintraub et al. 2009), including the Mini-Mental State Examination

(MMSE) (Folstein et al. 1975); the Logical Memory I subtest of the Memory-Scale-Revised (WMS-R) consisting of only one paragraph with 25 pieces of information to be recalled at immediate and delayed intervals (Wechsler 1987); Digit Span Forward and Backward on a range 0–12; Category Fluency for animals and vegetables; and the Digit Symbol subtest of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) that consists of at most 93 symbols to be replaced in at most 90 s (Wechsel 1981). The Visit 2 protocol repeated the above tests, but the Number Span Test was used in place of the Digit Span and Category Fluency for vegetables was omitted. The results of Number Span Test and Digit Span Test were harmonized to the same Digit Spans before analysis. The range of each test is described in Table 1. Six of these eight tests were included in the analysis: the MMSE was not included because it did not show sufficient variability at younger ages, while the Category Fluency for vegetable was not included because it was measured only once.

Statistical analysis

Test results that were judged to be invalid secondary to poor hearing or vision, environmental distractions, experimenter errors, or other physical limitations were not included in the statistical analysis. Patients' characteristics were summarized using mean and standard deviation or median and range (Table 1).

Group Discovery

We analyzed the data of each test independently using a Bayesian model-based clustering (BMBC) that we implemented in 2 steps. In step 1, we estimated one BMBC for each number of clusters (or “groups”) ranging between 2 and 20, and in step 2, we estimated the Bayesian information criterion (BIC) for each BMBC model and selected the model with minimum BIC (Fraley and Raftery 2002; Wiecki et al. 2015). For each pre-specified number of clusters, we modeled the repeated measure j of the test of an individual i in cluster w_k as a linear function of age at enrollment ($\text{age}.1_i$), change of age during the follow-up time for people who had data in both visits (d.age), sex, education (educ), and an indicator variable to denote whether the subject was a member of a longevous family or was a spouse control (spouse), using cluster specific effects, as shown in this equation:

$$y_{ij} \sim N(\mu_{ij}, \tau_y); i = 1, \dots, N, j = 1, 2$$

$$\begin{aligned} \mu_{ij} = & b0[w_k] + b.\text{age}[w_k]*\text{age}.1_i \\ & + b.\text{dage}[w_k]*\text{d.age}_{ij} + b.\text{sex}[w_k]*\text{sex}_i \\ & + b.\text{spouse}[w_k]*\text{spouse}_i \\ & + b.\text{educ}[w_k]*\text{educ}_i \end{aligned}$$

Note that this model can include individuals in the study with either one measurement at enrollment ($j = 1$) or with two measurements ($j = 1, 2$), one at enrollment and one at the second in-person visit. We modeled the probability of being a member of a cluster w_k using a multinomial distribution, with probability vector θ , where θ_k is the probability of cluster w_k in the population. The parameter $b.\text{age}[w_k]$ represents the cluster-specific, cross-sectional effect of age, while the parameter $b.\text{dage}[w_k]$ is the cluster-specific, longitudinal effect of age. We assumed normal distributions for the data of all the six neuropsychological tests that was confirmed by visual inspection of the data. We employed uninformative conjugate normal prior distributions for the regression parameters with 0 mean and variance 10, while we used a uniform Dirichlet prior for the parameters of the multinomial distribution that represents the latent group membership. Uniform priors were used to avoid biasing the results. We used Markov Chain Monte Carlo (MCMC) methods to compute Bayesian posterior estimates of the regression coefficients and of the parameters of the latent variables and assigned individuals to the group with maximum posterior probability. The MCMC for each analysis was run for at least 10,000 iterations, and final estimates were computed using a 10th of the simulated samples that was sufficient to reduce autocorrelation of parameters. We include the code specification of the analysis as [supplementary material](#). Figure 1 displays the most prevalent longitudinal patterns of change discovered by this analysis for the 6 neuropsychological tests, and the full set of results are shown in Supplement Figs. 1A–6A.

Annotation of clusters

We annotated the clusters with > 35 individuals discovered in the analysis of each of the six

Table 1 Summary of participants' characteristics by generation and visit

		Baseline	Visit 2
Proband generation	No participants (%)	1706 (34%)	438 (17%)
	No spouse (%)	205 (12%)	61 (14%)
	Age	91 (49–110)	96 (58–111)
	Birth year	1917 (1896–1958)	1921 (1904–1926)
	Sex (female)	953 (56%)	264 (60%)
	Education (years)	10 (1–17)	10 (2–17)
	DSST	29 (0–73)	27 (0–73)
	Animal fluency	14 (1–34)	14 (0–30)
	Logical Memory (Imd)	8 (0–21)	9 (0–22)
	Logical Memory (Dly)	6 (0–21)	6 (0–20)
	Digit Span Forward	7 (1–12)	7 (2–12)
	Digit Span Backward	5 (1–12)	5 (1–12)
Offspring generation	No participants (%)	3358 (66%)	2101 (83%)
	No spouse (%)	881 (26%)	495 (24%)
	Age	61 (25–93)	70 (45–98)
	Birth year	1947 (1919–1983)	1947 (1919–1972)
	Sex (female)	1853 (55%)	1148 (55%)
	Education (years)	14 (0–17)	14 (2–17)
	DSST	50 (0–93)	48 (0–90)
	Animal fluency	22 (5–45)	22 (0–43)
	Logical Memory (Imd)	13 (0–25)	14 (0–25)
	Logical Memory (Dly)	12 (0–25)	13 (0–25)
	Digit Span Forward	7 (2–12)	7 (1–12)
	Digit Span Backward	6 (1–12)	6 (1–12)

Median and range are reported for continuous variables. One hundred fifty-one individuals (3%) in the cohort were enrolled at the time of the second in-person visit and have only one measurement

neuropsychological tests by the rate of change that was calculated as the average of the absolute difference between scores at visit 1 and visit 2, relative to the score at visit 1. We estimated the association between the clusters and these aging markers that were measured at enrollment in LLFS participants:

- Risk factors: age, alcohol use, BMI, current smoker, familial longevity, pulse rate, systolic and diastolic blood pressure, years of education, weight, medications for hypertension, T2D, heart disease, and high cholesterol
- Self-reported medical history: atrial fibrillation, cancer, cataracts, congestive heart failure, COPD, coronary artery bypass grafting, deep vein thrombosis, dementia or Alzheimer's disease, depression, glaucoma, hypertension, macular degeneration, myocardial infarction, osteoporosis, Parkinson's disease,

stroke, TIA, T2D, thyroid disease, and valve replacement

- Markers of physical and pulmonary functions: grip strength, gait speed, FEV1
- Cognition as measured by the remaining 5 neuropsychological tests and the Mini-Mental State Examination (MMSE)
- 32 circulating biomarkers that correlate with age as described in (Sebastiani et al. 2016)

We also correlated clusters of each neuropsychological test with incident events of the same disease list during the follow-up, with changes of the other five neuropsychological tests and MMSE and with changes of physical and pulmonary function. These analyses used either age or sex-adjusted logistic regression or linear regression with random effects to model within family

Fig. 1 Main clusters discovered using model-based cluster analysis of each neuropsychological test. Single dots represent test scores at baseline, and segments represent repeated test scores at baseline and visit 2. Colors represent different clusters of individuals. For example, the green segments and dots of DSST scores represent individual assigned to cluster 2 of DSST. Clusters are sorted from the slowest to the fastest changers based on RCC: relative change of cognition that is defined for each test as absolute

difference between mean test score at visit 2 and mean test score at visit 1 relative to the mean test score at visit 1. Note that the cluster numbers represent the different grouping of individuals returned by the clustering algorithm in each neuropsychological test, and cluster labels of different tests are unrelated: for example, cluster 2 of DSST does not include the same people in cluster 2 of logical memory, immediate recall

correlation. Correlation with genotypes of *APOE* were tested using χ^2 test for genotype association, or genotype group association in which genotype groups were defined as E2 = E2/E2 or E2/E3, E3 = E3/E3, and E4 = E2/E4, E3/E4 or E4/E4.

To correct for multiple testing, we used a 5% FDR based on Benjamini correction (Benjamini and Hochberg 1995). The significant association between various factors and clusters of cognitive change were summarized by the circular graph in Fig. 2 that was generated with the *R* package

circize version 0.4.3 (Gu et al. 2014). The complete set of results is in Supplementary Tables 2–7. We annotated clusters with $N \geq 35$ individuals to have at least 80% power to detect standardized difference between clusters of at least 1, assuming a significance level of 0.5 E-02.

Agreement of different clusters and summary score

We estimated the pairwise agreement between the clusters discovered in the analysis of all six

Fig. 2 Significant associations between clusters of changes of the neuropsychological tests and patients' characteristics. Each connection represents a significant association, and the width is proportional to $-\log_{10}(p \text{ value})$. Dark blue, clusters of DSST change; dark orange, clusters of logical memory-immediate recall test change (Log Mem IMD); blue, clusters of logical memory-delayed recall test change (Log Mem DLY); orange, clusters of Digit Span Forward Test change (Digits Forward); light blue,

clusters of Digit Span Backward Test change (Digits Back); light orange, clusters of animal fluency test change (animal fl). Participants' characteristics are grouped as biomarkers (molecular and genetic markers), non-genetic risk factors and medications, disease history at enrollment, incident disease between enrollment and visit 2, physical functions at enrollment, and changes of physical and cognitive functions between enrollment and visit 2

neuropsychological tests using Jaccard index that compares agreement of stratification between two sets of different number of clusters (Birks 1987; Sebastiani and Perls 2016). To have an overall summary of the agreement between the 6 sets of clusters, we also generated a summary score by assigning points to cluster membership ranging from 0 for clusters with the fastest rate of cognitive decline to a maximum (equal to the

number of groups discovered for each test) for the slowest rate of cognitive decline (see Fig. 3). We analyzed the association between this score and biomarker signatures of aging based on 19 circulating biomarkers (Sebastiani et al. 2017), using linear regression adjusted for age at enrollment and the indicator variable of familial longevity (=member of family selected for familial longevity or spouse).

Table 2 Regression coefficients of model-based clustering of DSST change

DSST	Effects	Cluster 2	Cluster 8	Cluster 6	Cluster 5
Logical memory (immediate recall)	RCC ¹	0.11	0.12	0.15	0.19
	Age (C) ²	-0.65 (-0.72; -0.60)	-0.67 (-0.70; -0.64)	-0.66 (-0.68; -0.63)	-0.73 (-0.80; -0.68)
	Age (L) ³	-0.58 (-0.71; -0.46)	-0.48 (-0.55; -0.42)	-0.49 (-0.56; -0.42)	-0.42 (-0.58; -0.26)
	Sex (female)	5.86 (4.62; 7.18)	5.22 (4.49; 5.96)	4.87 (4.15; 5.60)	3.34 (1.92; 4.71)
	Spouse	-0.11 (-1.42; 1.19)	-0.55 (-1.37; 0.28)	-0.52 (-1.38; 0.34)	0.40 (-1.06; 2.13)
	Education (Y)	1.67 (1.51; 1.84)	1.43 (1.32; 1.55)	1.30 (1.19; 1.42)	1.26 (1.04; 1.46)
		Cluster 3	Cluster 1	Cluster 4	Cluster 2
		0.18 (0.17)	0.20 (0.2)	0.36 (0.54)	0.48 (0.44)
		-0.10 (-0.12; -0.07)	-0.13 (-0.14; -0.11)	-0.15 (-0.16; -0.14)	-0.14 (-0.16; -0.11)
		0.17 (0.10; 0.24)	0.05 (0.001; 0.09)	0.14 (0.09; 0.18)	-0.04 (-0.16; 0.05)
Logical memory (delayed recall)	Age (C) ²	0.48 (-0.08; 1.00)	1.01 (0.65; 1.35)	0.71 (0.37; 1.03)	0.69 (0.07; 1.02)
	Sex (female)	0.76 (0.15; 1.42)	0.16 (-0.25; 0.59)	0.20 (-0.18; 0.57)	-0.80 (-1.60; 0.005)
	Spouse	0.13 (0.06; 0.20)	0.29 (0.16; 0.38)	0.30 (0.25; 0.35)	0.28 (0.16; 0.39)
	Education (Y)		Cluster 5	Cluster 3	Cluster 2
		0.18 (0.16)	0.27 (0.3)	0.46 (0.61)	0.52 (0.29)
		-0.12 (-0.14; -0.10)	-0.14 (-0.16; -0.11)	-0.19 (-0.21; -0.17)	-0.21 (-0.29; 0.5)
		0.11 (0.05; 0.17)	0.10 (0.05; 0.15)	0.08 (-0.03; 0.15)	-0.86 (-1.31; -0.48)
		0.67 (0.64; 1.12)	1.11 (0.7; 1.67)	0.93 (0.48; 1.34)	0.24 (-1.65; 2.1)
		0.31 (-0.23; 0.87)	0.44 (-0.14; 0.91)	0.06 (-0.42; 0.58)	-0.38 (-2.47; 1.87)
		0.15 (0.08; 0.21)	0.35 (0.28; 0.47)	0.25 (0.17; 0.31)	0.25 (0.00; 0.55)
Animal fluency		Cluster 4	Cluster 8	Cluster 3	Cluster 7
		0.15 (0.13)	0.19 (0.17)	0.22 (0.19)	0.32 (0.34)
		-0.30 (-0.33; -0.27)	-0.24 (-0.26; -0.22)	-0.24 (-0.26; -0.22)	-0.13 (-0.18; -0.06)
		-0.05 (-0.14; 0.03)	0.00 (-0.05; 0.05)	-0.06 (-0.11; -0.00)	-0.15 (-0.36; 0.05)
		-0.24 (-0.94; 0.42)	-0.10 (-0.52; 0.34)	0.04 (-0.43; 0.48)	-0.37 (-1.78; 1.14)
		0.50 (-0.29; 1.37)	-4.11 (-5.10; -3.20)	5.54 (4.46; 6.25)	-0.38 (-2.08; 1.26)
		0.29 (0.19; 0.38)	0.31 (0.23; 0.38)	0.22 (0.14; 0.29)	0.38 (0.15; 0.61)
		Cluster 7	Cluster 3	Cluster 1	Cluster 5
		0.15 (0.12)	0.23 (0.22)	0.29 (0.13)	0.31 (0.32)
		-0.03 (-0.04; 0.00)	-0.03 (-0.03; -0.02)	-0.06 (-0.09; -0.03)	-0.03 (-0.31; -0.21)
Digits Backward		-0.13 (-0.20; -0.02)	0.04 (0.01; -0.07)	-0.20 (-0.35; -0.06)	-0.02 (-0.04; 0.00)
		0.25 (-0.21; 0.85)	0.08 (-0.12; 0.28)	0.27 (-0.44; 0.94)	0.11 (-0.02; 0.25)
		-0.05 (-0.52; 0.59)	-0.42 (-0.68; 0.00)	-0.25 (-0.31; 0.69)	0.01 (-0.25; 0.20)
		0.22 (0.10; 0.28)	0.20 (0.16; 0.24)	0.29 (0.23; 0.37)	0.11 (0.08; 0.13)
		Cluster 1	Cluster 4	Cluster 3	Cluster 2
		0.1 (0.09)	0.21 (0.13)	0.16 (0.15)	0.21 (0.15)
		-0.02 (-0.03; -0.01)	-0.03 (-0.04; -0.01)	0.01 (-2.8; 3.3)	-0.02 (-0.03; -0.01)
		-0.06 (-0.08; -0.03)	-0.20 (-0.28; -0.12)	0.03 (-2.97; 3.60)	-0.15 (-0.17; -0.12)
		0.00 (-0.18; 0.20)	-0.26 (-0.49; -0.01)	-0.07 (-3.15; 3.30)	0.06 (-0.12; 0.24)
		-0.11 (-0.33; 0.13)	-0.59 (-1.15; -0.21)	0.24 (-3.58; 2.92)	-0.46 (-0.70; -0.24)
Digits Forward		0.15 (0.12; 0.19)	0.25 (-0.21; 0.29)	0.25 (-2.76; 3.45)	0.19 (0.16; 0.22)
		Cluster 5			
		0.28 (0.21)			
		-0.02 (-0.03; -0.01)			
		-0.11 (-0.13; -0.07)			
		0.04 (-0.17; 0.26)			
		0.00 (-0.3; -0.30)			
		0.00 (-0.3; -0.30)			
		0.10 (-0.15; 0.18)			

¹ RCC = relative cognitive change abs (mean test at visit 2- mean test at visit 1)/mean test at visit 1² Regression coefficient and 95% credible interval for age at visit 1 (cross-sectional effect)³ Regression coefficient and 95% credible interval for time after visit 1 (longitudinal effect)

Clusters are sorted from slowest to fastest changers based on the RCC. For comparison, the column mixed-effect reports the estimate of the regression coefficients from the mixed-effect model

clusters of the six tests. We next report the major findings for each test.

Processing speed (DSST) The analysis discovered nine clusters of change in the Digit Symbol Substitution Test (DSST) score (Fig. S1A). Only four clusters included > 35 participants, and they are displayed in Fig. 1, sorted from the slowest to the fastest changers based on the relative cognitive change in Table 2. These clusters include individuals with average DSST score at enrollment ranging from 62.4 to 35.9 (Fig. S1B) and rate of change varying between -0.42 and -0.58 points per year. Table 2 reports both cross-sectional and longitudinal age effects of the models associated with the 4 largest clusters. The magnitude of the longitudinal effect in the BMBC analysis was slightly reduced compared with the cross-sectional age effect. For example, in cluster 2, the estimated decrease of DSST for each year of age at enrollment was 0.65 points, while the estimated decrease of DSST for each year in the study was 0.58 points. This difference could be the result of mortality of the oldest generation that left behind a cohort of younger and possibly more cognitively intact survivors. Sex and education were significant covariates in the BMBC model, and females appeared to have higher scores than males in all groups. Figure S1B includes the 26 variables associated with clusters of DSST change at FDR < 5%. Compared with participants in the other clusters, slowest changers (cluster 2) were older (mean age at enrollment 66.0 years), had the lowest level of education (11.4 years), and scored on average better in DSST at enrollment and in all the other neuropsychological tests, while fastest changers (cluster 5) scored on average worse in all the other tests at enrollment. Gait speed at enrollment was significantly different in participants with different patterns of DSST change, and slowest changers had on average the fastest gait speed, while fastest changers had the slowest gait speed. Five circulating biomarkers are associated with patterns of DSST change, and slowest changers had lower than average IL6 and NTproBNP, while fastest changers had lower than average albumin and hemoglobin levels and higher than average red cell distribution width at enrollment. Smoking, alcohol use, and higher pulse pressure were also associated with the clusters of DSST patterns and were predictive of faster decline. Interestingly, decreasing prevalence of cholesterol lowering medications predicted faster rate of decline. History of dementia, depression, Parkinson's, stroke, and TIA were associated

with patterns of DSST change, and participants with the fastest rate of DSST change (cluster 5) had the highest prevalence of depression at enrollment (21%), stroke (5%), dementia (4%), and incident dementia (9%). Clusters of DSST change differed also by the prevalence of *APOE* genotypes ($p = 0.017$), and slowest changers were enriched for E2 compared with all other groups (20% versus 15%) and had the lowest prevalence of E4 (17%), while fastest changers had a higher prevalence of the deleterious E4 genotype group (24%) (Table 3). Changes in Digit Span Forward Score were significantly associated with clusters of DSST change. Table S2 includes the complete results of the analysis.

Episodic memory (immediate and delayed recall) The analysis discovered four clusters of immediate recall change and five clusters of delayed recall change (Fig. 1). Both sets included a small cluster of fast changers with a low score at enrollment who declined more rapidly than other participants and a small cluster of slow changers with high score at enrollment who maintained higher score over time (Figs. S2–S3). The parameters of the cluster-specific models are in Table 2 and show that older age at enrollment was associated with lower score in all clusters but consistently with a recall effect noted in other studies (Salthouse 2019); increasing follow-up time was associated with improvement in both memory scores in almost all clusters: exceptions were the clusters of fastest decliners. Females scored higher than males in all groups, and the positive effect of education was consistent across groups. Faster gait speed was associated with slower change of immediate recall but not delayed recall (Figs. S2B, S3B). Slowest changers of immediate recall were on average older (66 years), had better memory at enrollment, had lower than average creatinine level but higher than average glucose level and absolute eosinophil counts, and had no history of dementia. The slowest changers of delayed recall had the lowest average body mass index (BMI, 26.5). Fastest changers of immediate recall were younger (64 years), had higher than average creatinine and glucose levels, and higher prevalence of E4 (29%). Fastest changes of delayed recall had the highest average BMI (27.7) (Tables S3 and S4). We also found a significantly higher prevalence of *APOE* E4 in the fastest changers of delayed recall (Table 3).

Semantic fluency After removal of very small clusters, the analysis of semantic fluency changes identified two

Table 3 Association of APOE with clusters of change patterns

Test	Cluster	APOE genotype groups ¹			p^2
		E2	E3	E4	
DSST	2	0.202	0.624	0.174	0.017
	8	0.144	0.665	0.192	
	6	0.150	0.650	0.200	
	5	0.156	0.605	0.239	
Logical memory (immediate recall)	3	0.154	0.640	0.206	0.007
	1	0.154	0.649	0.197	
	4	0.158	0.667	0.175	
	2	0.123	0.592	0.285	
Logical memory (delayed recall)	1 2 3 5	0.153	0.657	0.190	0.026
	4	0.178	0.574	0.248	
Digit Span (backward)	7	0.172	0.677	0.151	0.078
	3	0.156	0.662	0.182	
	1	0.116	0.643	0.241	
	5	0.158	0.634	0.208	

¹ E2 = E2/E2 or E2/E3; E3 = E3/E3 and E4 = E2/E4, E3/E4 or E4/E4

² p value from χ^2 test

Clusters sorted from slowest to fastest changers

small groups of faster and slower changers that included 20% of participants and two larger groups of more intermediate changers with 79% of participants (Fig. 1, Fig. S4A, Table 2). In all clusters, age at enrollment was associated with a worst score, but there was no significant effect of the increased age at follow-up (Table 2). History of dementia at enrollment, gait speed, and incident dementia were associated with patterns of this score change. The fastest changers (cluster 7) had the lowest average gait speed (0.89 m/s) and the largest prevalence (2%) of dementia at enrollment compared with the other groups. The incidence of dementia was also high (6%) but comparable with cluster 3 (Tables S5).

Attention/working memory (Digit Span Forward and Backward) The analysis of the Digit Span Forward Test identified five groups with more uniform numbers of subjects. One small group included participants with test scores that did not change over time (Table 2). Patients in different clusters had significantly different grip strength and gait speed, and the fastest changers had better grip strength and gait speed at enrollment compared with slowest changers ($p = 0.00013$ and 0.0014 , Table S6). In addition, decreasing prevalence of treatment for T2D or high cholesterol predicted faster rate of decline. The analysis of Digit Span Backward Test

identified four groups that weakly correlated with changes of other tests (Table S7). The cluster of fastest changers (cluster 5) was enriched for carriers of APOE E4, although the association did not reach statistical significance (Table 3, $p = 0.078$). In both tests, we noticed that several study participants reached a perfect score even at older ages. This ceiling effect is well known, and approximately 10% of the population across all ages can accurately complete both tests (Wechsel 2008).

Cluster agreement and summary score The Jaccard index showed a similarity between clusters of different tests in the 37–50% range. Clusters of changes in Digit Span Backward and delayed memory were the least similar, while clusters of changes in immediate and delayed memory tests were the most similar (Fig. S7). In order to summarize the similarity between the clusters of all neuropsychological tests, rather than the similarity of pairs of tests, we next generated a summary of cluster results by assigning increasing points to clusters that represent decreasing rate of change (Fig. 3) and by summing the points associated with each cluster across the six tests for each participant. The score was symmetrically distributed around 10, suggesting that the majority of individuals scored between 1 and 2 across

all tests, i.e., average decline across all domains. Less than 25% of participants scored < 8, which would happen when the majority of cognitive functions are in rapid decline, and less than 25% of participants scored > 12, which would happen when most cognitive functions are in slow decline. We also correlated this summary score with biomarker signatures of aging based on 19 circulating biomarkers (Sebastiani et al. 2017). This analysis showed that a signature of healthy aging characterized by lower than average inflammation was associated with a larger than average score, or slower decline, while other signatures characterized by increased inflammation, anemia, and worse kidney function were associated with lower than average score, or faster decline (Fig. 3c–f).

Discussion

Overview We discovered clusters of LLFS participants with statistically different patterns of age-related changes in multiple domains of cognitive function. We correlated the clusters with molecular markers, non-genetic risk factors, medical history, and with changes of other aging markers. This exploratory analysis showed that changes of different types of cognitive function are correlated with each other and with other markers of aging, including physical functions. Some factors, such as education, appear to affect changes in cognitive function similarly across many domains, whereas other factors including inflammatory biomarkers and genetic markers correlate with changes in specific cognitive domains.

Patterns For each of the six neuropsychological tests, we identified 4 or 5 main patterns of age-related change. With the exception of the two Digit Span Tests, in all other tests, we found small clusters of participants with either very rapid or very slow decline. These patterns of change would be missed by a regression-type analysis that describes the overall age trend of the 6 tests or analyses that examine only the cross-sectional changes of score or only the longitudinal changes of score. For example, clusters 2 and 6 of DSST have similar cross-sectional age effects but different longitudinal age effects, while clusters 4 and 3 of animal fluency have different cross-sectional age effects but similar

longitudinal age effects (Table 2). Interestingly, extreme patterns of decline were typically seen only in one test and rapid decline in multiple cognitive domains, or simultaneous maintenance of very good cognitive function in many domains was not common in this cohort (Newman et al. 2011). The patterns of change were characterized by both substantial variabilities of the scores at the beginning of the study and varying rates of change over time. We did not observe scores that start at a similar level in younger people and decline at a different rate as people age, and the wide variability in test scores at age 40–50 is consistent with the theory that cognitive decline starts early in life (Harada et al. 2013). These results suggest a correlation between rates of change of various domains of cognitive function and baseline values. Additional longitudinal assessment of cognitive state of LLFS participants to be administered between 2020 and 2023 will provide data to investigate these relations more in detail.

In addition, we found a large number of participants with excellent cognitive functions at very old age. For example, the cluster of slowest changers of DSST included participants with an average score around 60 at age 65: about 10 points higher than the normative value reported for white Americans of the same age and with a college degree (Schneider et al. 2015). Virtually, all clusters included a mix of younger and older individuals with the same pattern of change of neuropsychological test score (Table S8), consistent with evidence that cognitive age may be more dependent upon health and function than chronological age (Wahlin et al. 2006). Some of the clusters of slow changers included members of the proband generation, confirming that factors associated with extreme longevity may be important for preservation of cognitive function (Andersen et al. 2019b). More in-depth analyses of such clusters could identify neuroprotective factors that promote longevity.

Risk factors Patterns of change in each test were associated with morbidity and molecular and environmental risk factors (Fig. 2). The analysis confirmed the effects of age, sex, and education on global cognitive decline (Schneider et al. 2015) with size and direction of effects similar to those reported in Weintraub's study of cognitively normal older adults (Weintraub et al. 2009). The associations between gait speed at enrollment and patterns of change in all tests were also consistent with previous findings (Buracchio et al. 2010; Rosano et al.

2008), and our analysis confirmed that faster gait speed correlates with slower decline in all domains. Our analysis also identified risk factors that are specifically associated with some and not all patterns of change: for example, the analysis suggests that low inflammation is associated with slower change of processing speed, while poor kidney functions or increased anemia predicts faster declines of memory (Singh-Manoux et al. 2014).

Consistent with the hypothesis that different pathways may lead to specific types of cognitive decline, our analysis discovered risk factors of changes in specific cognitive domains that may provide new ways to monitor and predict cognitive decline in older adults more accurately. For example, while changes in gait speed or grip strength may not be sufficient to differentially predict changes in processing speed rather than attention, our analysis suggests that history of neurological disease and levels of IL6 and NTproBNP may be useful to predict changes in processing speed, while history of T2D may be more specific for predicting changes in attention.

Interestingly, we found an *APOE* effect only on changes of three tests, and our analysis confirms the deleterious effect of *APOE* E4 on memory and the lack of effect of *APOE* on semantic fluency (Salmon et al. 2013). Moreover, the increased prevalence of *APOE* E2 in the cluster of slowest changers of DSST is consistent with the growing evidence of a neuroprotective role of *APOE* E2 (Kim et al. 2017) that could be linked to the longevity-associated mechanism of E2 (Sebastiani et al. 2018). A comprehensive analysis of genetic variants associated with patterns of change of various cognitive functions is beyond the scope of this manuscript, but consistently with the large heritability of results on neuropsychological tests found by us and other investigators (Barral et al. 2014; Matteini et al. 2010), we expect that genetic factors in addition to *APOE* may be an important determinants of these changes.

Our analysis showed that treatments for T2D and high cholesterol were associated with patterns of slower decline of DSST and Digit Span Tests, but did not appear to correlate with patterns of change in other cognitive domains. Post-marketing analyses have suggested that the use of statin may be associated with reversible cognitive impairment, but there is substantial controversy on this topic (Schultz et al. 2018). Our

analyses suggest that statin and common treatments for T2D like metformin may actually have a beneficial effect on processing speed. Given the clinical implications of this finding, the results need to be replicated in multiple studies.

Agreement between tests The analysis of the six tests showed only moderate agreement between changes of different cognitive functions. The relatively low agreement and the specificity of some of the risk factors speak to the importance of measuring a wide spectrum of cognitive functions and the need for multiple risk factors to monitor the cognitive health of aging individuals. Some overlapping associations between clusters and risk factors may be due to the fact that cognitive tests are multidimensional and not able to measure an isolated cognitive process. For example, an individual with reduced attentional capacity will perform more poorly on tests of episodic memory due to inattention rather than memory impairment. Similarly, changes in specific cognitive functions, such as processing speed (Lindenberger and Ghisletta 2009), or non-cognitive factors, such as hearing and vision impairments, may affect a wide array of cognitive functions (Anstey et al. 2003). The lack of strong agreement in the results however suggests that the cluster analysis detected changes in specific cognitive processes measured by the individual tests. For example, clusters that show large change in episodic memory but less change in attention are more likely to reveal associations with risk factors that are specific to memory impairments. From these clusters with isolated impairment, we can learn more about the risk factors associated with pathological changes due to a specific disease process, for example, vascular disease versus Alzheimer's disease.

Limitations Our results are currently limited to LLFS participants, and this work will need to be replicated in independent cohorts to show general applicability. We provide code and details of the statistical analysis to support independent replication from other investigators. In addition, the availability of at most 2 repeated measurements of the cognitive tests in LLFS participants limited the type of analysis to linear models of age, while more complex trajectories could be captured with longer longitudinal measurements as recently shown (Salthouse 2019). Duncan and Duncan (Duncan and Duncan 2009) warned about the limitation of using growth curve modeling with

only two time points and emphasized the need to interpret the results with caution. LLFS participants will be administered another cognitive assessment between 2020 and 2023, and these data will be useful to expand the analysis to include more complex and possibly more specific trajectories of change. A related issue is the substantial rate of loss to follow-up that is due to a combination of high mortality of the oldest generation (~75%) and dropout in the younger generation (~35%). We analyzed the effect of dropout rate in the estimation of the effect of familial longevity on the rate of cognitive decline, and our analyses suggest a pattern of non-informative dropout (Andersen et al. 2019a). However, the mortality in the older generation is likely to reduce the longitudinal effect of age on cognitive decline at the oldest ages, and we need better methods of statistical analysis to approach this problem. We did not adjust the model-based cluster analyses for within-family correlation, and more advanced statistical analyses could include random effects to address this limitation and reduce false positive associations. All our analyses were adjusted by sex, but it is reasonable to expect that there may be sex-specific patterns of cognitive decline determined by exposure to different hormones (Matyi et al. 2019; Núñez et al. 2019). Additional analyses in larger data sets are warranted to better characterize the role of sex and particularly sex hormones on the decline of cognitive function.

Conclusions

This analysis discovered patterns of changes of neuropsychological tests that share some common themes. The correlation between patterns of changes of cognitive functions in multiple domains and patterns of changes of physical functions is consistent with the work of other investigators that aging affects many domains simultaneously (Clouston et al. 2013). However, our analyses also show that patterns of changes of neuropsychological tests have some unique features and risk factors that could be useful to predict cognitive aging more specifically and are consistent with the theory of multiple pathways of cognitive aging. In order to better characterize these different pathways, it will be important

to discover the dynamic relationships between changes in different aging domains.

Acknowledgments PS: designed the study, conducted the analyses, and drafted the manuscript; MD and BS: statistical analyses; SA: study design and interpretation of results; BT, SC, NS, KC, and TTP: interpretation of the results. All authors edited the manuscript.

Funding information This work was supported by the National Institute on Aging (NIA cooperative agreements U01-AG023712, U01-AG23744, U01-AG023746, U01-AG023749, U01-AG023755, U19AG023122, P30AG031679, R01AG061844, R21AG056630, and K01-AG057798), the National Institute of General Medical (NIGMS) Interdisciplinary Training Grant for Biostatisticians [T32 GM74905], and the Paulette and Marty Samowitz Family Foundation (TP).

Compliance with ethical standards

Competing interests The authors declare that they have no conflicts of interest.

References

- Andersen SL, Du M, Cosentino S, Schupf N, Rosso A, Perls TT, Sebastiani P (2019a) Slower decline in processing speed is associated with familial longevity in the long life family study J Gerontol, B in review.
- Andersen SL, Sebastiani P, Dworkis DA, Feldman L, Perls TT. Health span approximates life span among many supercentenarians: compression of morbidity at the approximate limit of life span. J Gerontol A Biol Sci Med Sci. 2012;67:395–405. <https://doi.org/10.1093/gerona/gle223>.
- Andersen SL, Sweigart B, Sebastiani P, Drury J, Sidlowski S, Perls TT. Reduced prevalence and incidence of cognitive impairment among centenarian offspring. J Gerontol A Biol Sci Med Sci. 2019b;74:108–13. <https://doi.org/10.1093/gerona/gly141>.
- Anstey KJ, Hofer SM, Luszcz MA. A latent growth curve analysis of late-life sensory and cognitive function over 8 years: evidence for specific and common factors underlying change. Psychol Aging. 2003;18:714–26. <https://doi.org/10.1037/0882-7974.18.4.714>.
- Bae HT, Sebastiani P, Sun JX, Andersen SL, Daw EW, Terracciano A, et al. Genome-wide association study of personality traits in the long life family study. Front Genet. 2013;4:65. <https://doi.org/10.3389/fgene.2013.00065>.
- Barral S, et al. Common genetic variants on 6q24 associated with exceptional episodic memory performance in the elderly. JAMA Neurol. 2014;71:1514–9. <https://doi.org/10.1001/jamaneurol.2014.1663>.

- Barral S, et al. Exceptional memory performance in the long life family study. *Neurobiol Aging*. 2013;34:2445–8. <https://doi.org/10.1016/j.neurobiolaging.2013.05.002>.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J Roy Stat Soc B Met*. 1995;57:289–300.
- Birks HJB. Recent methodological developments in quantitative descriptive biogeography. *Ann Zool Fenn*. 1987;24:165–78.
- Blanken AE, Jang JY, Ho JK, Edmonds EC, Han SD, Bangen KJ, et al. Distilling heterogeneity of mild cognitive impairment in the National Alzheimer Coordinating Center database using latent profile analysis. *JAMA Netw Open*. 2020;3:e200413. <https://doi.org/10.1001/jamanetworkopen.2020.0413>.
- Buracchio T, Dodge HH, Howieson D, Wasserman D, Kaye J. The trajectory of gait speed preceding mild cognitive impairment. *Arch Neurol*. 2010;67:980–6. <https://doi.org/10.1001/archneurol.2010.159>.
- Clouston SA, et al. The dynamic relationship between physical function and cognition in longitudinal aging cohorts. *Epidemiol Rev*. 2013;35:33–50. <https://doi.org/10.1093/epirev/mxs004>.
- Cloutier S, Chertkow H, Kergoat M-J, Gauthier S, Belleville S. Patterns of cognitive decline prior to dementia in persons with mild cognitive impairment. *J Alzheimers Dis*. 2015;47:901–13. <https://doi.org/10.3233/JAD-142910>.
- Cosentino S, Schupf N, Christensen K, Andersen SL, Newman A, Mayeux R. Reduced prevalence of cognitive impairment in families with exceptional longevity. *JAMA Neurol*. 2013;70:867–74. <https://doi.org/10.1001/jamaneurol.2013.1959>.
- Duncan TE, Duncan SC. The ABC's of LGM: an introductory guide to latent variable growth curve modeling. *Soc Personal Psychol Compass*. 2009;3:979–91. <https://doi.org/10.1111/j.1751-9004.2009.00224.x>.
- Elo IT, Mykita L, Sebastiani P, Christensen K, Glynn NW, Perls T. Age validation in the long life family study through a linkage to early-life census records. *J Gerontol B Psychol Sci Soc Sci*. 2013;68:580–5. <https://doi.org/10.1093/geronb/gbt033>.
- Folstein MF, Folstein SE, McHugh PR. Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189–98.
- Fraley C, Raftery AE. Model-based clustering, discriminant analysis, and density estimation. *J Am Stat Assoc*. 2002;97:611–31.
- Goh JO, An Y, Resnick SM. Differential trajectories of age-related changes in components of executive and memory processes. *Psychol Aging*. 2012;27:707–19. <https://doi.org/10.1037/a0026715>.
- Gu Z, Gu L, Eils R, Schlesner M, Brors B. Circlize implements and enhances circular visualization in R. *Bioinformatics*. 2014;30:2811–2. <https://doi.org/10.1093/bioinformatics/btu393>.
- Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. *Clin Geriatr Med*. 2013;29:737–52. <https://doi.org/10.1016/j.cger.2013.07.002>.
- Kim YJ, et al. Protective effects of APOE e2 against disease progression in subcortical vascular mild cognitive impairment patients: a three-year longitudinal study. *Sci Rep*. 2017;7:1910. <https://doi.org/10.1038/s41598-017-02046-y>.
- Lee JS, Cho SK, Kim HJ, Kim YJ, Park KC, Lockhart SN, et al. Prediction models of cognitive trajectories in patients with nonamnestic mild cognitive impairment. *Sci Rep*. 2018;8:10468. <https://doi.org/10.1038/s41598-018-28881-1>.
- Lindenberger U, Ghisletta P. Cognitive and sensory declines in old age: gauging the evidence for a common cause. *Psychol Aging*. 2009;24:1–16. <https://doi.org/10.1037/a0014986>.
- Lovden M, Bergman L, Adolffson R, Lindenberger U, Nilsson LG. Studying individual aging in an interindividual context: typical paths of age-related, dementia-related, and mortality-related cognitive development in old age. *Psychol Aging*. 2005;20:303–16. <https://doi.org/10.1037/0882-7974.20.2.303>.
- Lunn D (2013) The BUGS book: a practical introduction to Bayesian analysis. *Texts Stat Sci Series*.
- Marron MM, et al. Heterogeneity of healthy aging: comparing long-lived families across five healthy aging phenotypes of blood pressure, memory, pulmonary function, grip strength, and metabolism. *Geroscience*. 2019. <https://doi.org/10.1007/s11357-019-00086-y>.
- Matteini AM, et al. Heritability estimates of endophenotypes of long and health life: the long life family study. *J Gerontol A Biol Sci Med Sci*. 2010;65:1375–9. <https://doi.org/10.1093/gerona/glq154>.
- Matyi JM, Rattinger GB, Schwartz S, Buhusi M, Tschanz JT. Lifetime estrogen exposure and cognition in late life: the Cache County Study. *Menopause*. 2019;26:1366–74. <https://doi.org/10.1097/gme.0000000000001405>.
- Murman DL. The impact of age on cognition. *Semin Hear*. 2015;36:111–21. <https://doi.org/10.1055/s-0035-1555115>.
- Newman AB, et al. Health and function of participants in the Long Life Family Study: a comparison with other. *Cohorts Aging* (Albany NY). 2011;3:63–76.
- Núñez F, Maraver MJ, Colzato LS. Sex hormones as cognitive enhancers? *J Cogn Enhancement*. 2019;4(228):233. <https://doi.org/10.1007/s41465-019-00156-1>.
- Rosano C, Newman AB, Katz R, Hirsch CH, Kuller LH. Association between lower digit symbol substitution test score and slower gait and greater risk of mortality and of developing incident disability in well-functioning older adults. *J Am Geriatr Soc*. 2008;56:1618–25. <https://doi.org/10.1111/j.1532-5415.2008.01856.x>.
- Salmon DP, et al. Age and apolipoprotein E genotype influence rate of cognitive decline in nondemented elderly. *Neuropsychology*. 2013;27:391–401. <https://doi.org/10.1037/a0032707>.
- Salthouse TA. Trajectories of normal cognitive aging. *Psychol Aging*. 2019;34:17–24. <https://doi.org/10.1037/pag0000288>.
- Schneider AL, et al. Normative data for 8 neuropsychological tests in older blacks and whites from the atherosclerosis risk in communities (ARIC) study. *Alzheimer Dis Assoc Disord*. 2015;29:32–44. <https://doi.org/10.1097/WAD.0000000000000042>.
- Schultz BG, Patten DK, Berlau DJ. The role of statins in both cognitive impairment and protection against dementia: a tale of two mechanisms. *Transl Neurodegener*. 2018;7:5. <https://doi.org/10.1186/s40035-018-0110-3>.
- Sebastiani P, et al. APOE alleles and extreme human longevity. *J Gerontol A Biol Sci Med Sci*. 2018. <https://doi.org/10.1093/gerona/gly174>.
- Sebastiani P, Hadley EC, Province M, Christensen K, Rossi W, Perls TT, et al. A family longevity selection score: ranking sibships by their longevity, size, and availability for study.

- Am J Epidemiol. 2009;170:1555–62. <https://doi.org/10.1093/aje/kwp309>.
- Sebastiani P, Perls TT (2016) Detection of significant groups in hierarchical clustering by resampling. *Front Genet* 7 doi: <https://doi.org/10.3389/fgene.2016.00144>
- Sebastiani P, Sun FX, Andersen SL, Lee JH, Wojczynski MK, Sanders JL, et al. Families enriched for exceptional longevity also have increased health-span: findings from the Long Life Family Study. *Front Public Health*. 2013;1:38. <https://doi.org/10.3389/fpubh.2013.00038>.
- Sebastiani P, Thyagarajan B, Sun F, Honig LS, Schupf N, Cosentino S, et al. Age and sex distributions of age-related biomarker values in healthy older adults from the long life family study. *J Am Geriatr Soc*. 2016;64:e189–94. <https://doi.org/10.1111/jgs.14522>.
- Sebastiani P, Thyagarajan B, Sun F, Schupf N, Newman AB, Montano M, et al. Biomarker signatures of aging. *Aging Cell*. 2017;16:329–38. <https://doi.org/10.1111/accel.12557>.
- Singh-Manoux A, Dugravot A, Brunner E, Kumari M, Shipley M, Elbaz A, et al. Interleukin-6 and C-reactive protein as predictors of cognitive decline in late midlife. *Neurology*. 2014;83:486–93. <https://doi.org/10.1212/WNL.0000000000000665>.
- Small BJ, Dixon RA, McArdle JJ. Tracking cognition-health changes from 55 to 95 years of age. *J Gerontol B Psychol Sci Soc Sci*. 2011;66(Suppl 1):i153–61. <https://doi.org/10.1093/geronb/gbq093>.
- Wahlin A, MacDonald SW, deFrias CM, Nilsson LG, Dixon RA. How do health and biological age influence chronological age and sex differences in cognitive aging: moderating, mediating, or both? *Psychol Aging*. 2006;21:318–32. <https://doi.org/10.1037/0882-7974.21.2.318>.
- Wechsler D (1981) Wechsler adult intelligence scale – revised. Harcourt brace Jovanovich [for] psychological Corp; revised edition.
- Wechsler D. Wechsler adult intelligence scale—fourth edition administration and scoring manual. San Antonio: Pearson; 2008.
- Wechsler D. WMS-R: Wechsler memory scale-revised manual. Harcourt Brace Jovanovich: Psychological Corp; 1987.
- Weintraub S, et al. The Alzheimer’s disease centers’ uniform data set (UDS): the neuropsychologic test battery. *Alzheimer Dis Assoc Disord*. 2009;23:91–101. <https://doi.org/10.1097/WAD.0b013e318191c7dd>.
- Wiecki TV, Poland G, Frank MJ. Model-based cognitive neuroscience approaches to computational psychiatry: clustering and classification. *Clin Psychol Sci*. 2015;3:378–99.
- Wilson RS, Beckett LA, Barnes LL, Schneider JA, Bach J, Evans DA, et al. Individual differences in rates of change in cognitive abilities of older persons. *Psychol Aging*. 2002;17:179–93.

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