Brain Structure and Function (2020) 225:519–536
https://doi.org/10.1007/s00429-019-02005-5

ORIGINAL ARTICLE

Brain aging and psychometric intelligence: a longitudinal study

L. Jäncke[1,2,3] · S. Sele[1,2] · F. Liem[2] · J. Oschwald[2] · S. Merillat[2]

Received: 20 April 2019 / Accepted: 6 December 2019 / Published online: 20 December 2019

Abstract

In this study, we examined a large sample of 231 generally healthy older adults across 4 years with regard to several brain anatomical measures (volumes of total grey matter volume: GM, normal appearing white matter: NAWM, lateral ventricle: LV, and white matter hypointensities: WMH) and psychometric intelligence (verbal and non-verbal). The dataset comprised four measurement occasions (baseline, 1-, 2-, and 4-year follow-ups). With this longitudinal data set, we evaluated level–level, level–change, and change–change relationships between the anatomical and psychometric measures using latent growth curve models. Our analyses indicate that GM and NAWM decreased significantly over the course of 4 years with annual percent changes of −0.73% and −0.79%, respectively. WMH and LV volumes increase with annual percent changes of 7.3% and 4%, respectively. Verbal and nonverbal IQ measures remained stable in our sample. In addition, we uncovered evidence for level–level and -change associations between several of the brain anatomical measures. With regard to brain-IQ associations, we observed a positive level–level association for GM and NAWM, indicating that participants with larger brain volumes demonstrate higher IQ measures. No substantial evidence was identified for level- or change–change associations between any of the brain metrics and the IQ measures. Taken together, these results suggest that while healthy older adults demonstrated age-related neuroanatomical decline over a time span of 4 years, these degenerative changes are not necessarily linked to simultaneous cognitive deterioration.

Keywords Brain aging · Psychometric intelligence · Longitudinal study · Level–level-association · Level–change-association · Change–change-association

Introduction

It is a well-known fact that aging is accompanied by substantial neuroanatomical and neurophysiological changes. For example, total brain volume and volumes of several subareas decrease while volumes of white matter hyperintensities and ventricles increase with age, mainly after the age of 55 (Fjell and Walhovd 2010). In terms of neurophysiological measures, blood flow, power in all EEG frequency bands, and resting-state functional connectivities decrease with aging (Sleimen-Malkoun et al 2015; Miraglia et al 2016; Catchlove et al 2018; Bangen et al 2018). Correlated with these neuroanatomical and neurophysiological changes, performance in most cognitive functions decline (Salthouse 2010), supporting the idea that the neuroanatomical, neurophysiological, and cognitive changes are correlated or even depend on each other. However, although the above-mentioned findings are well agreed upon, one has to keep in mind that the majority of the studies reporting and supporting these findings are based on cross-sectional data.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00429-019-02005-5) contains supplementary material, which is available to authorized users.

✉ L. Jäncke
lutz.jaencke@uzh.ch

S. Sele
silvano.sele@uzh.ch

F. Liem
franziskus.liem@uzh.ch

J. Oschwald
jessica.oschwald@gmail.com

S. Merillat
susan.merillat@uzh.ch

[1] Department of Neuropsychology, Psychological Institute, University of Zurich, Zürich, Switzerland

[2] University Research Priority Program (URPP), "Dynamic of Healthy Aging", University of Zurich, Zürich, Switzerland

[3] Division Neuropsychology, Institute of Psychology, University of Zurich, Binzmühlestrasse 14/25, 8050 Zurich, Switzerland

Cross-sectional designs can generate data more rapidly. However, in contrast to longitudinal studies, they are associated with a number of disadvantages (e.g., impossibility to analyze intraindividual time courses and confounding with cohort effects). As pointed out in our recent literature review regarding correlated changes in brain structure and cognitive ability (Oschwald et al 2019a), longitudinal studies are urgently required to advance our understanding of the associations between these two domains in the context of healthy aging.

Depending on the available data and the research question of interest, one can examine level–change and/or correlated change associations. Level–change relationships describe any association between a cross-sectional measure (e.g., structural brain properties or cognitive performance at the baseline of a study) and longitudinal changes in the respective other domain. For example, individuals with more intact structural brain features may show reduced age-related cognitive decline over a follow-up period compared to individuals with a lower degree of healthy brain tissue. On the other hand, higher levels of cognitive ability might also serve protectively from premature age-related brain degeneration. Finally, it is also plausible that bidirectional relationships between these two domains occur. However, because level–change relationships are only quasi-longitudinal, it is difficult to disentangle whether these associations reflect within-person change processes or rather cross-sectional differences between individuals. If longitudinal data are available for both domains, it is possible to investigate change–change relationships. Such correlated change associations can occur simultaneously or in a lagged fashion, such that previous changes in one domain relate to subsequent changes in the other domain. Simultaneous correlated changes between structural brain features and cognitive ability provide insights into cross-domain associations that occur within the same time frame, reflecting either direct change–change relationships between the two domains or the influence of a third variable on both change trajectories.

Several longitudinal studies have been published so far examining neuroanatomical and cognitive changes longitudinally across healthy adults of different age ranges (Schmidt et al 1999, 2005; Cohen et al 2001; Raz et al 2007, 2008; Silbert et al 2008; Leow et al 2009; Charlton et al 2010; Grimm et al 2012; Persson et al 2012, 2016; Fjell et al 2014, 2016, 2017; Lövdén et al 2014; Mak et al 2015b, 2015a; Daugherty et al 2015; Ritchie et al 2015a, 2015c; Bender et al 2016; Möller et al 2016; Köhncke et al 2016; Gorbach et al 2017; Sala-Llonch et al 2017; Leong et al 2017; Hohman et al 2017; Moon et al 2017; Anblagan et al 2018; Song et al 2018; Yuan et al 2018).

While most of these studies have demonstrated age-related anatomical changes, they differ in many aspects hindering to draw a final conclusion. For example, sample size, cultural background of the studied subjects, education, health status, age-ranges, used anatomical measures, number of time points, and statistical models vary substantially between these studies. Only a few works have been published so far examining longitudinal associations between neuroanatomical and cognitive changes in subjects exclusively older than 65 years (Silbert et al 2008; Leow et al 2009; Lövdén et al 2014; Mak et al 2015b, 2015a; Ritchie et al 2015a, 2015c; Köhncke et al 2016; Sala-Llonch et al 2017; Hohman et al 2017; Moon et al 2017; Anblagan et al 2018).

Eight of these studies have examined level–change associations (Silbert et al 2008; Leow et al 2009; Mak et al 2015b; Ritchie et al 2015a, c; Hohman et al 2017; Moon et al 2017; Anblagan et al 2018) with five of them reporting significant level–change associations, however, between different cognitive functions and brain measures (Ritchie et al 2015a, c; Hohman et al 2017; Moon et al 2017; Anblagan et al 2018). Significant change–change associations have been reported by eight studies (Silbert et al 2008; Lövdén et al 2014; Ritchie et al 2015a, c; Köhncke et al 2016; Sala-Llonch et al 2017; Moon et al 2017; Anblagan et al 2018).

Given the heterogeneity of findings for the different cognitive domains and brain areas, the present work was designed to examine the structure–cognition relationship between global measures of psychometric intelligence and several global brain metrics during the course of a longitudinal measurement. The dataset used for analysis comprises psychometric and MRI data collected from a large sample of healthy and non-demented older subjects, examined four times within a period of approximately 4 years. Thus, we are in the position to model the longitudinal time course over this relatively short period more reliably than using only two-time points as has been done in most studies of this type. Only two of the above-mentioned studies with subjects of comparable age included more than two-time points for both a measure of brain structure and cognitive ability (Silbert et al 2008; Hohman et al 2017).

In addition, incorporating four occasions allows us to estimate and compare both linear and nonlinear trajectories (King et al 2018). We also believe that it is helpful and even necessary to independently replicate and validate findings from different groups working with various MRI scanners, using different psychological tests, and working with subjects with varying cultural backgrounds. In this project, we will, therefore, replicate the research of the above-mentioned research groups [particularly the study of Ritchie et al. (2015c), which longitudinally examined a large sample of subjects in a similar age range and using similar brain anatomical measures as we will use in this paper] but with a dataset comprising four-time points obtained during a period of 4 years. In addition, we will also work with intracranial volume (ICV) as a covariate (which was not used in the

Ritchie et al. study) to correct for general head size differences. This approach will also correct for sex/gender differences, which should essentially disappear or at least diminish substantially (Mathalon et al 1993; Jäncke et al 2015).

In this study, we focus on the structure–function relationship between several global brain metrics and psychometric intelligence. Different from the aforementioned studies, we will use a German variant of the Thurstone intelligence test, allowing us to work with a general IQ estimate and six primary factors representing different psychological functions (verbal comprehension, spatial processing, memory functions, reasoning, word fluency, and perceptual speed). In our study, we aimed to work on the following issues:

1. Modeling the time course of brain atrophy for global anatomical measures (volumes of total gray matter, normal-appearing white matter, white matter hypointensities, and lateral ventricle) and cognitive performance changes (verbal and non-verbal IQ) (denoted as Model 1 in our analysis).
2. Examining the level–change and change–change associations between the time courses of brain atrophies for the different brain measures (Model 2).
3. Examining the level–change and change–change associations between the time courses of verbal and non-verbal IQ (Model 3).
4. Modeling the level–change and change–change associations between the time courses of cognitive functions (for the two IQ measures) and brain atrophies (for the different brain measures) (Models 4 and 5).

Methods

Subjects

Longitudinal cognitive and MRI data were taken from the Longitudinal Healthy Aging Brain (LHAB) database—an ongoing project conducted at the University Research Priority Program (URPP) 'Dynamics of Healthy Aging' of the University of Zurich (Zöllig et al 2011). We used data from four measurement occasions (baseline: tp1, 1-year follow-up: tp2, 2-year follow-up: tp3, 4-year follow-up: tp5) (see Fig. 1). The baseline dataset included 231 participants (mean age = 70.84 ∓ 5.08). During the following years, the number of subjects dropped to 210 at tp2 (mean age 72.0 ∓ 5.2), to 197 at tp3 (mean age: 72.7 ∓ 4.8) and to 173 at tp5 (mean age: 74.23 ∓ 4.27). 24 participants of the baseline sample took part in an additional 3-year follow-up (tp4). For several subjects, we did not have both the anatomical and all the IQ measures at each time point. For the anatomical measures, the number of subjects was 231 at tp1, 207 at tp2, 196 at tp3, 24 at tp4, and 166 at tp5. For non-verbal-IQ, the number of subjects was 227 at tp1, 207 at tp2, 190 at tp3, 164 at tp5. For the verbal-IQ, the number of subjects was 230 at tp1, 206 at tp2, 190 at tp3 and 163 at tp5. The proportion of females was 49% at baseline, 48% at 1-year follow-up, 46% at 2-year follow-up and 46% at 4-year follow-up. At each measurement occasion, participants completed an extensive battery of neuropsychological and psychometric cognitive tests and underwent brain imaging. Brain imaging was conducted at the University Hospital of Zurich, either after cognitive testing or on a separate day within a span of 1–2 weeks of the cognitive assessment.

Eligibility criteria for study participation were age ≥ 64, a score of ≥ 26 on the Mini-Mental State Examination [MMSE; (Folstein et al 1975)], right-handedness, fluent German language proficiency, no self-report of any neurological or psychiatric disease, and no contraindications to MRI. The self-reported physical and mental health of the sample at baseline, as measured by the SF-12 (Ware et al 1996a) was 50.7 ± 7.4 and 54.6 ± 6.4, respectively, which, according to the SF-12 norms indicates above-average health (Ware et al 1996b). Although those subjective health indicators slightly decline across time, they still remain above average. With respect to diabetes and hypertension, conditions that are common in older adults, the prevalence in our sample is consistent with or below the values reported on population-level for older adults. The study was approved by the ethical committee of the canton of Zurich. Participation was voluntary and all participants gave written informed consent in accordance with the declaration of Helsinki. The data of this sample has been used in previous publications of our group (Madhyastha et al 2014; Liem et al 2015; Hirsiger et al 2016; Valizadeh et al 2017, 2018).

For the present analyses, structural MRI and IQ data were used. To avoid bias due to selective exclusion of participants with incomplete data, we use the full sample (data present for MRI measures and/or IQ measures) in this article. The IQ scores of one subject at the 4-year follow up were excluded because of a large drop in language IQ score from 135 at the 2-year follow up to 41 at the 4-year follow up. This subject further only received a score of 22 at the 4-year follow up in the MMSE. Some more subjects ($N = 8$) received an MMSE score < 26 in follow-up measurements but were not excluded because they showed an MMSE score > 26 in later follow-up measurements ($N = 4$) or showed no unusual drop in IQ scores ($N = 4$). All remaining subjects at the 1-year, the 2-year, the 3-year and the 4-year follow up were still healthy and non-demented according to the criteria applied at wave 1.

Measuring psychometric intelligence

Psychometric intelligence was measured with a German intelligence test (Leistungsprüfsystem: IQ), which is based

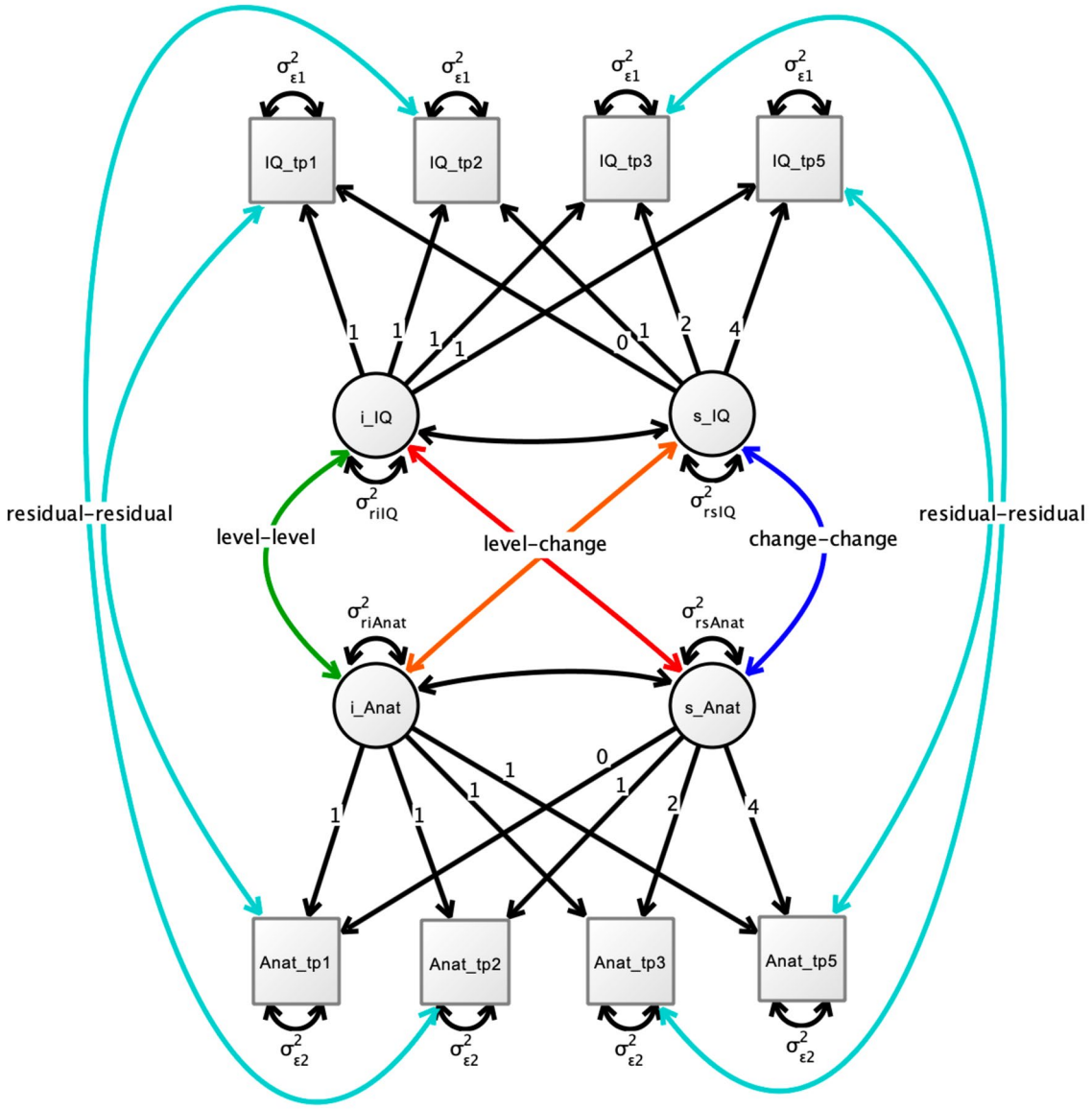

Fig. 1 Simplified schematic latent growth curve diagram associating the trajectories of an intelligence test measure (IQ) to the trajectories of an anatomical measure (Anat) over four-time points (tp1, tp2, tp3, tp5). The diagram shows a bivariate random slope model for the anatomical and the IQ measure. The circles represent the latent variables (random intercepts with prefix *i* and random slopes with prefix *s*), the squares represent the observed variables. One-headed arrows represent regression paths with fixed loadings on observed variables, two-headed arrows represent variances and covariances. The colored covariances provide information about the association between the IQ and the anatomical measure. As in all fitted models, the residuals were assumed to be the same over the time points and the residual-residual associations were also assumed to be the same for each time point. For simplicity, the covariates (age at study entry and total intracranial volume) were omitted in the diagram. Further not shown is the training effect slope for the IQ measures and the quadratic slope for the normal appearing white matter. For the anatomical measures, there were additionally some participants with five-time points

on Thurstone's intelligence model (Thurstone 1938). Here we used the version for subjects older than 50 years up to 90 years (Sturm et al 1993). The split-half reliability of this test ranges between $r = 0.89$ and $r = 0.97$ depending on the particular subtest. Using the norms provided by the test battery, the mean IQ was 120.6 (sd = 6.7) for the total sample. Men and women significantly differed with males demonstrating a slightly larger IQ (121.7 ± 6.7) than women (119.4 ± 5.9) ($t(224) = 2.62$, $p = 0.0095$, Cohen's $d = 0.35$). IQ scores could be slightly overestimated since the subjects mostly fall exactly in the age range between two age categories (55–69 years and 70–90 years). We have used the normalization for the age category of 70–90 years for the entire sample. For the following statistical analyses, we use IQ raw scores. Since verbal and non-verbal cognitive measures often dissociate, we calculated a language IQ (IQ-L)

and a non-language IQ (IQ-nonL) score. The IQ-L measures comprise the subscales 1 + 2 (word recognition, general education, verbal knowledge, and spelling), 5 (word recognition, word fluency), 6 (word fluency and verbal knowledge), and 12 (word recognition) of the LPS50 + , while the IQ-nonL measure comprises the subscales 3 [detection of (un)regularities in geometric figures], 4 (like scale3 but with numbers and letters), 7 (recognition of regularities), 8 + 9 (spatial imagination, symbol comparison), 10 (recognition of the essential in a character, introversion, mental independence), 11 (recognition of fragmented images and words, visual memory perception), 13 (perception speed), and 14 (compare with test 13). Here, we computed separate IQ scores for language and non-language functions since it is known that language and non-language functions are controlled by mostly non-overlapping brain regions (Voyer 1996).

Image acquisition

MRI data were acquired with a 3.0 T Philips Ingenia scanner (Philips Medical Systems, Best, The Netherlands). We have described the image acquisition procedure in several of our recent papers, thus we partly reiterate what we have mentioned in these papers (Madhyastha et al 2014; Liem et al 2015; Hirsiger et al 2016; Valizadeh et al 2017, 2018). T1-weighted images were recorded with a gradient echo sequence (3D turbo field echo, 160 sagittal slices, slice thickness = 1 mm, in-plane resolution = 1 × 1 mm, FOV = 240 × 240 mm, repetition time = 8.18 ms, echo time = 3.80 ms, flip angle = 8°).

Image preprocessing

FreeSurfer (v5.3) as implemented in the FreeSurfer BIDS-App (Gorgolewski et al 2017) was used to obtain measurements of cortical and subcortical anatomy (Fischl et al 2002, 2004; Destrieux et al 2010). After completing the standard recon-all pipeline, measurements for cortical thickness, surface area, and volume were extracted for the regions of the Destrieux (aparc.a2009s) parcellation scheme (Destrieux et al 2010). Subcortical and global volume measurements were also extracted from FreeSurfer's *aseg* segmentation. To ensure independence between time points, FreeSurfer's longitudinal sectional analysis stream was used. For this paper, we estimated the following anatomical measures: (1) intracranial volume (ICV: EstimatedTotalIntraCranialVol), (2) total gray matter volume (GM: including lhCortex + rhCortex + SubCortGray + CerebellumGM), (3) total cerebral white matter volume (normal-appearing white matter without white matter hypointensity: NAWM), (4) white matter hypointensity volume (WMH: WM_hypointensities), and (5) lateral ventricle volume (LV). We have chosen these anatomical measures because they have been used in several papers examining the relationship between anatomical measures and psychometric intelligence (Luders et al 2007, 2008) and in recent papers examining age-related anatomical changes (Ritchie et al 2015a, b).

Statistical analysis

In this study, we examined several questions: first, we are interested to examine whether and how the neuroanatomical and IQ measures change over the course of 4 years. Second, we examined whether these changes are correlated (level–change and change–change associations). These associations are evaluated separately for the anatomical data (GM, NAWM, LV, and WMH) and for the IQ data. Finally, we were interested in the level–change and change–change associations across domains. Thus, we tested whether the anatomical measures at baseline are predictive for the IQ changes occurring over the course of the following years, or the other way around (whether the IQ measures at baseline are predictive for the later occurring anatomical changes) and if the changes in brain structure and cognitive ability are correlated For our statistical analysis, we used latent growth curve modeling (Grimm et al 2016) to describe the change of the neuroanatomical and the IQ measures. The advantage of this technique over other techniques is that one can model and statistically test change–change and level–change associations within one latent growth model. In addition, all the available information is exploited even in the presence of missing values, which for various reasons regularly occur in longitudinal studies. According to our study questions, we are interested in whether the latent growth variables of these measures and their residuals covary with each other. To achieve this, it is important to find a properly fitting model for each measure based on model fit statistics (e.g., likelihood ratio tests) (similar as in McArdle et al 2004).

Additionally, we are modeling the association between person-specific residuals per time point for each measure. At each time point, we obtain a residual term, which is the difference from the observed value to the person-specific expected value. It is important to note that these residuals do not only include random measurement errors, they may also include true deviations from the expected values of the fitted models. Therefore, it is reasonable to ask whether the residuals of different measures covary with each other at each time point. For example, will a person who scores higher than expected (person-specific) at baseline in measure 1 also score higher than expected at baseline in measure 2 (residual-residual association)? Our latent growth model also allows for modeling and testing these residuals and the associated covariance parameters in one model.

The estimated total intracranial volume (ICV) and the age of the participants at baseline were included as covariates

for the latent growth variables of the anatomical and the IQ measures in all fitted models. By including ICV as covariate, head size differences are controlled for (Mathalon et al 1993) and most sex differences are eliminated or at least diminished (Jäncke et al 2015). Both covariates were mean-centered.

Based on the individual models for each measure, we built a multivariate growth model associating the four anatomical measures to each other (model 1) and a bivariate growth model associating the language and non-language IQ scores to each other (model 2). For the next step of testing cross-domain associations, we decided to build 2 models, one linking GM and NAWM with the IQ measures (model 3) and one linking WMH and LV with the IQ measures (model 4). This was done because associating the language and non-language IQ scores to all anatomical measures within one model would have resulted in a model with too many free parameters. This model layout for the multidomain associations was chosen because for GM and NAWM positive associations with IQ are expected, while for WMH and LV the expected associations are negative.

To answer our main questions about possible evidence for associations between anatomical and IQ measures, we checked sequentially whether the inclusion of all the level–level associations, all the level–change associations, all the change–change associations, and all the residual-residual associations improved the model fit. This sequential procedure was started with a simple model containing none of these associations. The model fit improvement was measured with the likelihood ratio test. A p value < 0.05 was applied as an indicator for associations (or at least for one association). Given the exploratory nature of our analyses, we performed several model comparisons. We did not apply any statistical control for multiple comparisons over the sequential tests and over the fitted models. So far, no clear guidelines exist on how to best control for multiple comparisons in complex multivariate SEM models. Thus, most researchers using these models do not apply any Type I error control (Smith and Cribbie 2013). While one recommendation is to control for the number of hypothesis tests within an SEM model (Cribbie 2007), it is a subjective decision whether this number is based on the structural or on the measurement model.

For this group of parameters that do not improve model fit significantly, we report the parameter estimates from a model containing all the associations. The parameter estimates are reported in unstandardized form as covariances but also in standardized form as correlations for better readability. Of note, the z-statistic in the parameter tables refers to the covariance parameter estimates, but not to the correlation estimates, which would be slightly different. Simple symmetric confidence intervals for the correlation estimates are further reported. These may be smaller than -1 or bigger than 1 when there were estimation difficulties with the involved parameters, indicating very high uncertainty in these estimates.

The data of the multivariate growth models have been z-transformed or were left untransformed as appropriate when there were estimation difficulties with the covariance parameters. Missing values were accounted for using full maximum likelihood estimation, which results in unbiased estimates given that missing values are missing at random (Baraldi and Enders 2010). Because we were interested in the aging process of healthy people, this assumption seems appropriate. All analyses were performed in R (R Core Team 2014). The latent growth curve models were fitted with the *lavaan* package version 0.6–3 (Rosseel 2012). A simplified schematic representation of the latent growth model we used here is shown in Fig. 1.

Reliability of measures

The reliability of a measurement can be quantified with the intraclass correlation coefficient, which is the ratio of the between-subject variability to the total variability (the sum of the between and the within-subject variability) (Bartlett and Frost 2008). However, we were interested in the individual changes in the subjects. If we want high precision in our association estimates between the measures, it would be important that the within-subject variability is small compared to the person-specific yearly changes. The residual variances can be transformed to a repeatability coefficient (Bland and Altman 2003; Bartlett and Frost 2008), which quantifies below which value the absolute difference between two measurements on the same person will be with 95% probability, defined as $1.96 \times \sqrt{2} \times$ standard deviations of the residuals. Usually, a repeatability coefficient is calculated on the assumption that the measurement object does not change so that the only variance of measurement is originating from the measurement instrument itself. Nevertheless, comparing these approximate repeatability coefficients to the magnitude of the random slope variance and to the average slope of each measure provides some valuable information about the precision of each measure.

Results

Time course of the anatomical and psychological measures

Figure 2 shows the time course of the anatomical and psychological measures. In this figure the anatomical measures are corrected for ICV, thus minimizing (and partly eliminating sex differences). A detailed demonstration of the descriptive values is shown in the supplementary sTable2.

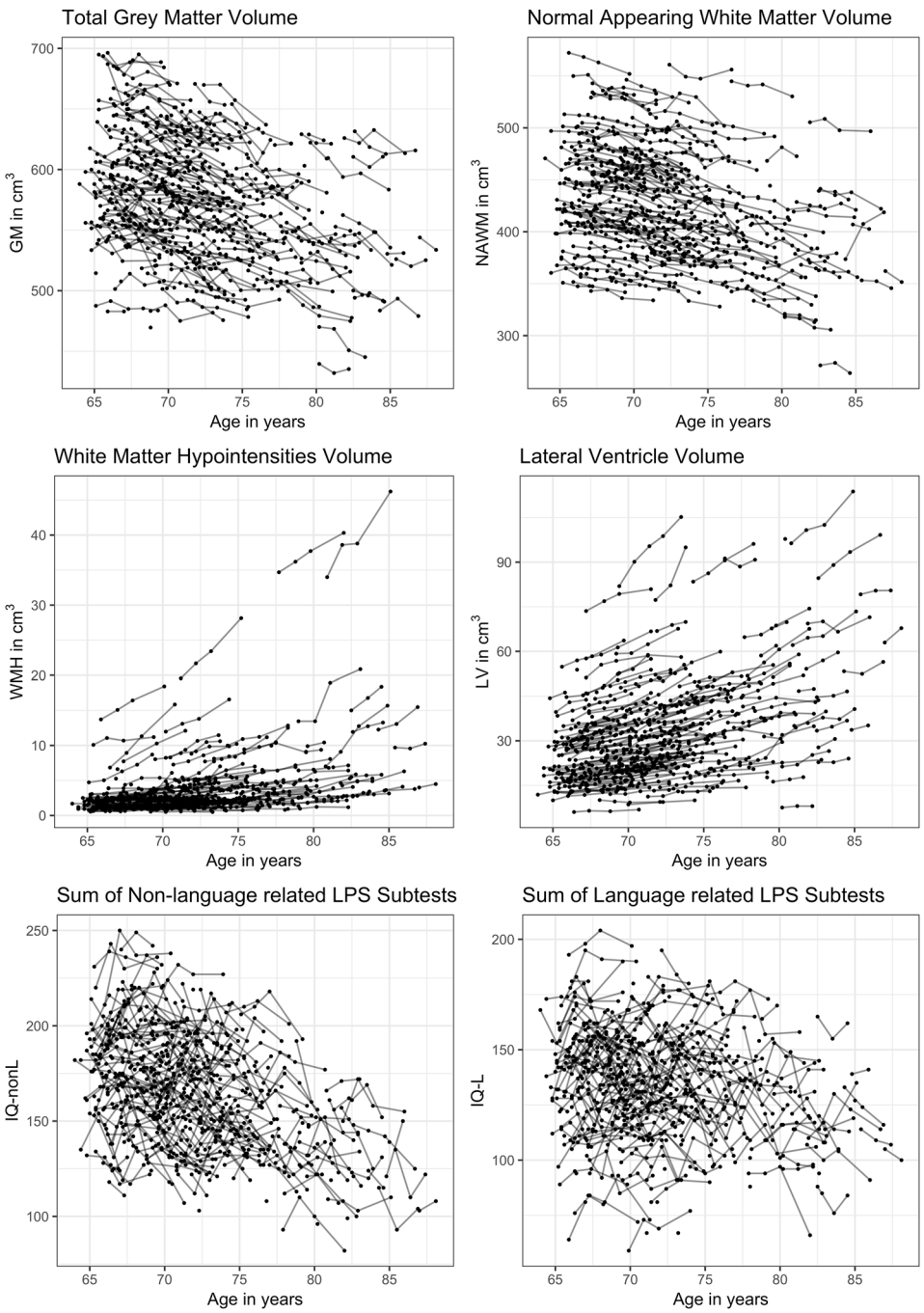

Fig. 2 Individual trajectories of the anatomical and IQ measures across age. WMH are shown as raw values. For statistical analysis of WMH, we used the log-transformed values

In Table 1, mean percent changes and Cohen's d effect size measures between two successive time points are shown. For GM the percent changes per year range between −0.43 and −1.3% with an average percent change per year of −0.74%. When demonstrating these changes as Cohen's effect size measures, the d range between $d=-0.42$ and $d=-0.63$ with an average $d=-0.46$. The mean percent changes for NAWM ranged between −0.43 and −0.91% with a mean percent change of −0.79%. In terms of effect size, these changes range between a $d=-0.32$ and $d=-1.04$, which amounts to a mean effect size per year of $d=-0.56$. Thus, the decreases of GM and NAWM are of medium size.

For WMH and LV substantial increases were observed between two successive time points. In terms of percent change per year for WMH, these values range between 6.35 and 7.7% with an average percent change per year of 7.4%. In terms of Cohen's d the effect sizes range between $d=0.43$ and 0.27 with an average $d=0.5$. The LV increases in terms of percent change per year measures

Table 1 Mean differences between two successive time points (tp) for all anatomical variables and the two IQ measures

Differences	Change in %	Mean (95% CI)	Cohen's *d*	*N*
GM tp2-tp1	−0.72	−4.30 (−5.40, −3.19)	−0.53	207
GM tp3-tp2	−0.59	−3.23 (−4.33, −2.13)	−0.42	190
GM tp4-tp3*	−0.43	−2.68 (−6.51, 1.16)	−0.29	24
GM tp5-tp4*	−1.30	−7.65 (−13.06, −2.24)	−0.63	22
GM tp5-tp3	−1.27	−7.49 (−8.85, −6.13)	−0.85	164
NAWM tp2-tp1	−0.58	−2.50 (−3.33, −1.67)	−0.41	207
NAWM tp3-tp2	−0.47	−1.92 (−2.78, −1.07)	−0.32	190
NAWM tp4-tp3*	−0.54	−2.18 (−4.24, −0.13)	−0.45	24
NAWM tp5-tp4*	−1.08	−4.68 (−6.68, −2.68)	−1.04	22
NAWM tp-tp3	−1.82	−7.88 (−8.96, −6.81)	−1.13	164
WMH tp2-tp1	6.45	0.26 (0.19, 0.34)	0.49	207
WMH tp3-tp2	6.35	0.28 (0.18, 0.37)	0.43	190
WMH tp4-tp3*	9.37	0.30 (0.15, 0.45)	0.86	24
WMH tp5-tp4*	7.72	0.20 (0.02, 0.39)	0.48	22
WMH tp5-tp3	13.97	0.52 (0.37, 0.66)	0.54	164
LV tp2-tp1	4.28	1.30 (1.11, 1.49)	0.94	207
LV tp3-tp2	3.65	1.16 (0.94, 1.39)	0.74	190
LV tp4-tp3*	4.01	1.13 (0.64, 1.62)	0.98	24
LV tp5-tp4*	5.17	1.69 (0.92, 2.45)	0.98	22
LV tp5-tp3	7.49	2.35 (2.03, 2.67)	1.13	164
IQ-non-L tp2-tp1	0.79	1.17 (−0.66, 3.00)	0.09	204
IQ-non-L tp3-tp2	1.10	1.76 (0.10, 3.43)	0.15	186
IQ-non-L tp5-tp3	−4.72	−7.65 (−9.38, −5.91)	−0.69	158
IQ-L tp2-tp1	1.77	1.90 (0.50, 3.31)	0.19	205
IQ-L tp3-tp2	1.14	1.72 (0.16, 3.27)	0.16	186
IQ-L tp5-tp3	−2.82	−4.34 (−6.47, −2.21)	−0.32	159

Shown are the mean percent changes (change in %), the mean (in cm^3) with the upper and lower bounds of the 95% confidence intervals, Cohen's *d* (based on variance of the differences between the time points), and the number of subjects; tp1: time point 1, tp2: time point 2, tp3: time point 3, tp4: time point 4, and tp5: time point 5

*Some subjects were measured on a yearly basis from time point 4 to time point 5, and time point 3 to 4

range between 4 and 5.1% with an average increase per year of 4.2%. The effect size range between $d = 0.74$ and $d = 0.98$ amounting to an average $d = 0.84$. Thus, the volume increases for WMH are of medium and those for LV are of large size.

Model 1: time-course analysis for anatomical and psychological measures

The model fit of the time courses revealed that all anatomical measures significantly changed over the 4 years. Figure 2 shows the individual trajectories of all measures over the time course of 4 years. The estimated parameters and the associated statistics of the time course models are shown in the Supplementary material on the sTable 3. For the sake of clarity, we only show the summary statistics of the time course analysis (Table 2). GM and NAWM revealed significant linear declines over the 4 years. For NAWM, we also identified a quadratic trend in addition to the linear trend, which was qualified by a steeper decrease from time point 3 to time point 5. For GM, the linear decline we identified no substantial between-subject variance (after adjusting for initial age and total intracranial volume). The NAWM decline was also largely similar across all subjects. This will be further evaluated when we report and discuss the change–change associations between GM and NAWM.

WMH and LV, on the other hand, demonstrated a significant linear increase over time. For WMH we used log-transformed volume measures to stabilize variances. For both of these measures, there was substantial between-subject variance in the estimated individual slopes.

For all four brain measures, baseline age was a strong predictor of the intercepts. In each model, it was of similar magnitude as the slope parameter. For LV and for the WMH, baseline age was also a significant predictor for the volume increase over time. Thus, older subjects demonstrated stronger changes in LV and WMH.

For both psychometric IQ measures (IQ language and IQ non-language), we observed nonlinear changes across the four-time points qualified by initial increases (until 2-year follow-up) and subsequent decreases (between 2-year and 4-year follow-up). Since re-test and habituation effects, which may cause performance stability or increases, are well-known issues in longitudinal studies (Salthouse 2010), we modeled the time course with a function representing training effects (0, 1, $\sqrt{2}$, 1) over the four-time points in addition to a linear slope. The underlying idea was to model the sample mean more appropriately. Instead of using a (0, 1, 1, 1) step function [as was done in (Yuan et al 2018)], we were assuming a nonlinear fixed training effect across time points. Thereby, we were expecting a slight increase of the training effect from the 1-year-follow-up to the 2-year-follow up, and afterward, a slight decrease of the training effect on the 4-year follow-up because there were 2 years between the 2-year and the 4-year follow up. We came up with the exact numbers of the training effect function by visually observing the patterns in the data and using model fit indices. Model fit clearly improved from $p < 0.0001$ to $p = 0.07$ for IQ-L and from $p < 0.0001$ to $p = 0.02$ for IQ-nonL by the inclusion of

Table 2 *p*-values obtained from the latent growth model fit testing for linear and quadratic trends of the anatomical and IQ measures

	GM	NAWM	WMH	LV	IQ-L	IQ-nonL
Intercept	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Slope	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Quadratic slope	–	< 0.0001	–	–	–	–
Age on intercept	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.006	< 0.0001
Age on slope	–	–	0.004	0.005	< 0.0001	0.002
Age on quadratic slope	–	–	–	–	–	–

The detailed results of these model tests are shown in the supplementary material on sTable 3 (– indicates $p > 0.05$)

this additional training regressor. Of note, the inclusion of this fixed average training effect had little influence on the main questions of our study, which were about the associations between the latent variables, because it reduced the within-subject error variance only by a small amount. The estimated parameters and the associated statistics of the time course models are shown in sTable 3 of the Supplementary material.

Using the parameter estimates of the random intercept and the residual variance reported in sTable 3 (neglecting the very small random slope variability), we obtained the following reliability estimates for the anatomical and the IQ measures: GM: 0.967, NAWM: 0.987, WMH: 0.990; LV: 0.995, IQ-nonL: 0.894, IQ-L: 0.910. For the anatomical and the IQ measures, we obtained relatively high repeatability coefficients [GM: 16.32 cm^3, WM: 12.16 cm^3, WMH (on log scale): 0.17, LV: 2.3 cm^3, IQ-nonL: 23.41, IQ-L: 19.83]. The results indicate that the LV was measured with the highest precision followed by the WMH. LV and WMH also had the smallest residual-to-random-slope variance ratios. The other anatomical and the IQ measures had quite high residual-to-random-slope variance ratios, indicating a larger uncertainty in the estimated individual slopes in these measures. A higher uncertainty in the person-specific slope estimates will result in a higher uncertainty concerning the level–change and change–change associations.

Model 2: level–level, level–change, change–change, residual–residual associations for anatomical measures

Associating the anatomical measures with each other in a multivariate growth curve model (Model 2) revealed evidence for several level–level, level–change, change–change and residual–residual relationships between the anatomical measures (see for a summary the sTable 4 shown in the Supplementary material). The significant level–level-associations are shown in Fig. 3. Clear evidence was found for all the level–level associations between the anatomical measures. Person-specific intercepts of the anatomical measures (GM and NAWM) were positively associated with each other and negatively associated with the person-specific intercepts of LV and WMH. The person-specific intercepts of the LV and the WMH were positively associated with each other. The largest correlation was found between the intercepts of GM and the intercepts of NAWM ($r = 0.56$). The other (absolute) correlations ranged between $r = 0.15$ and 0.23 (Fig. 3).

Evidence was found for several significant level–change associations between the anatomical measures with absolute correlations ranging between $r = 0.22$ and $r = 0.58$. Positive level–change associations were found between the intercept (level) and the slopes (change) of LV, between the intercept of NAWM and the slope of GM, as well as between the intercept of IV and the slope of WMH. Negative correlations were found between the intercept of GM and the slopes of LV and WMH. There were also negative correlations between the intercept of NAWM and the slopes of GM and LV. There were no further significant level–change associations.

There were also four significant change–change associations between the anatomical measures. It is clearly evident that there is a positive association between the slopes of LV and WMH ($r = 0.47$). Negative change–change associations were found between LV and GM, between LV and NAWM, as well as between WMH and GM. These correlation estimates were quite large in magnitude, ranging from -0.53 to -0.6. However, these estimated correlations involving the global anatomical slopes (GM and NAWM) have to be taken with care because they were estimated with relatively large uncertainty. No further substantial evidence for change–change associations between the anatomical variables was identified.

Clear evidence was found for all residual–residual associations between the anatomical measures with correlation magnitudes ranging in absolute values from 0.14 to 0.47. The residuals of the WMH and the LV were positively associated with each other. The residuals of the global anatomical measures (NAWM and GM) were negatively associated with the residuals of the LV. The residuals of the GM models were further negatively associated with the residuals of the WMH. However, the residuals of NAWM were positively

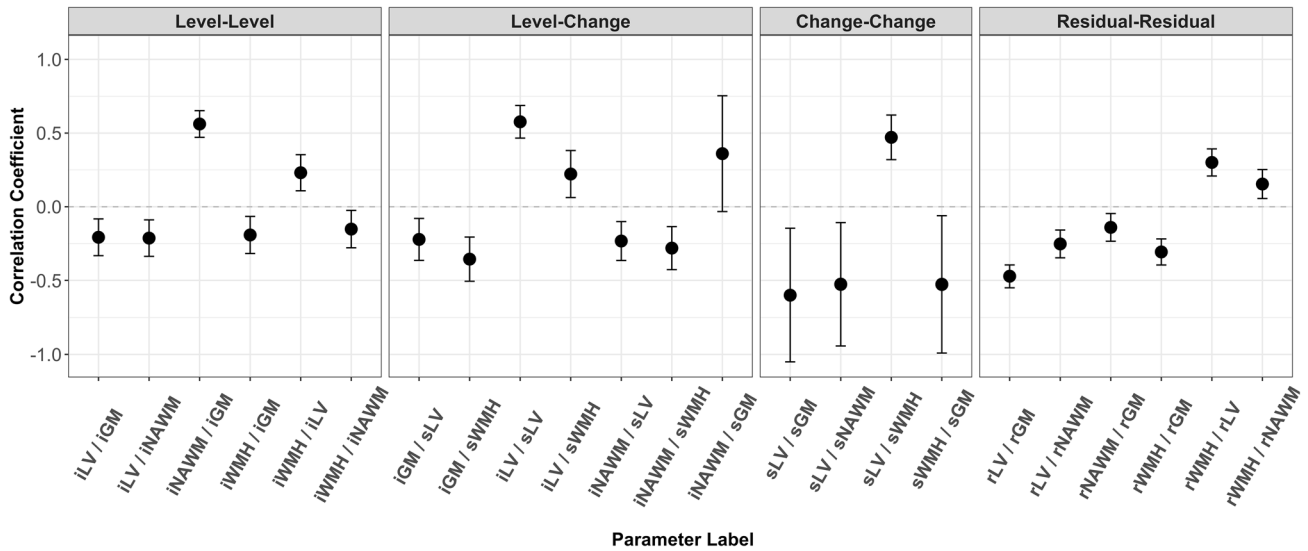

Fig. 3 Significant (at least $p < 0.05$) level–level-associations, level–change-associations, change–change-associations, and residual-residual-associations for the anatomical measures shown as correlation coefficients. Indicated are also the 95% confidence intervals for the correlation coefficients. The prefix "*i*" indicates statistical tests based on the intercepts (level). The prefix "*s*" indicates statistical tests based on the slopes (change). The prefix "*r*" indicates statistical tests based on the residuals. Detailed statistics (etc. estimates, *p* values, confidence intervals, etc.) for all variables and associations (even for those, which are not significant) are shown in the sTable 4 of the supplements

associated with the residuals of the WMH. The residuals between the global anatomical measures were negatively associated with each other.

Model 3: level–level, level–change, change–change, residual–residual associations for the IQ measures

Clear evidence for large positive level–level associations between the language and the non-language IQ scores were identified ($r = 0.66$). Evidence for a small positive association between the residuals of the language IQ scores and the residuals of the non-language IQ scores further was found ($r = 0.15$), indicating that if a person scores higher than expected in the language IQ tests, it is expected that this person also scores higher than expected in the non-language IQ tests at the same time point. No evidence for level–change or change–change associations between the language and the non-language IQ scores were found. However, the uncertainty in the level–change estimate was large and the uncertainty in the change–change estimate was very large.

Model 4 and Model 5: level–level, level–change, change–change, residual–residual associations for the IQ measures and the anatomical measures

For the sake of clarity and ease of interpretation, we calculated two growth models to associate the anatomical measures to the IQ measures. In the first model (Model 4) we used GM and NAWM and in the second model (Model 5) WMH and LV. Associating GM and NAWM to both IQ measures using sequential model comparisons revealed evidence for level–level relationships between the anatomical and the IQ measures (χ^2 (4) = 16.96, $p = 0.002$), but did not reveal evidence for level–change (χ^2 (8) = 10.08, $p = 0.26$), change–change (χ^2 (4) = 8.58, $p = 0.072$), or residual–residual associations (χ^2 (4) = 4.82, $p = 0.31$). The estimated covariance parameters are shown in the Supplementary sTable 6. Associating the LV and the WMH to the IQ measures in sequential model comparisons revealed no evidence for level–level (χ^2 (4) = 7.14, $p = 0.13$), level–change (χ^2 (8) = 11.38, $p = 0.18$), change–change (χ^2 (4) = 3.7, $p = 0.45$) or residual–residual associations (χ^2 (4) = 4.17, $p = 0.38$). The estimated covariance parameters are shown in the Supplementary sTable 7. These parameter estimates suggest that there may eventually be small negative level–level correlations between the LV and the IQ measures.

The intercepts of GM and NAWM were both positively associated with the intercepts of the non-language IQ scores. Evidence for a positive level–level association was further found between GM and the language IQ scores. The magnitude of these correlation estimates ranged between 0.18 and 0.25. Negative correlations ranging from −0.13 to −0.18 were found for the level–level associations between the LV and the language and non-language IQ scores and between the WMH and the non-language IQ scores. No further

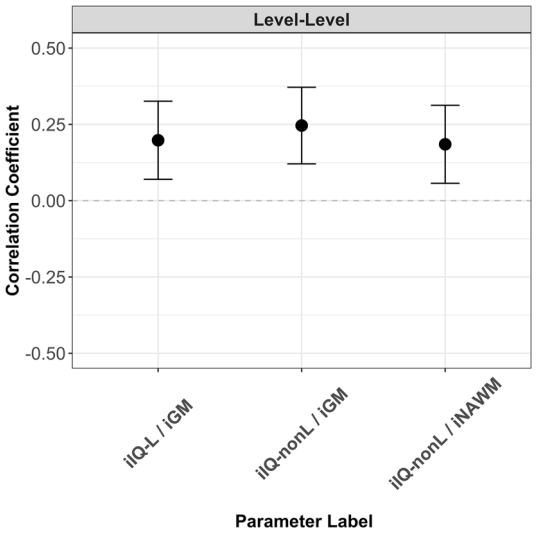

Fig. 4 Significant (at least $p<0.05$) level–level-associations for the IQ and brain measures indicated as correlations. The prefix "*i*" indicates statistical tests based on the intercepts (level)

level–level and level–change-associations were significant (Fig. 4).

Discussion

This paper was designed to examine the change of major brain anatomical measures (volumes of total gray matter, normal appearing white matter, lateral ventricle, and white matter hypointensity) and IQ measures in a relatively large cohort of healthy and cognitively well-functioning elderly subjects using a longitudinal study protocol. As a further study aim, we are interested in examining change–change associations between the different time courses. Here, we provide a focus on change–change associations between the different anatomical measures but also on change–change associations between the IQ and anatomical measures. Finally, we examined level–change associations for different brain and IQ measures. In the following sections, we will discuss our findings in the context of published results.

Longitudinal changes across the 4 years

For all anatomical measures, we identified significant changes over the 4 years. While NAWM and GM demonstrated significant linear declines, WMH and LV significantly increase over the 4 years. For NAWM, we identified a quadratic trend in addition to the linear trend, which was qualified by a steeper decrease from tp 3 to tp 5. The mean percent changes per year for GM and NAWM were −0.73% and −0.79%, respectively. The percent changes per year for WMH and LV were 7.3% and 4% per year, respectively. Thus, the changes are larger for WMH and LV than for GM and NAWM. Referring to effect sizes, which are more appropriate here because they take the sample variability into account, the changes are strongest for LV with large effect sizes, while the effect sizes for WMH, GM, and NAWM are moderate. Our results for loss of healthy GM and NAWM and increases of WMH and LV are in accordance with previous cross-sectional (Walhovd et al 2011) and longitudinal studies (for a summary see Oschwald et al 2019a).

There are, however, some differences between our findings and those reported in previous studies. First, the WMH increases are substantially smaller in our study (approximately 7%/year) than those reported by Ritchie and colleagues (2015c) (11%/year). These differences might be due to methodological differences. Notably, Ritchie et al. investigated WM hyperintensities using FLAIR images, while the present work investigated WM hypointensities on T1w images. In addition, the sample used by Ritchie et al. are also on average older than our sample. Secondly, Walhovd et al. (2011) reported in their review on cross-sectional studies stronger GM declines than those for WM. However, restricting their analysis to the age range covered in the present study (70–75 years), they found more WM loss. Ritchie et al. (2015c) also reported stronger WM than GM decline with percent decreases per year of 1% and 0.6%, respectively, results which they interpret as being consistent with the results reported by Walhovd et al. (2011). In our study, there was, in fact, no difference in terms of decline between GM and NAWM. The declines expressed as percent changes per year or as effect sizes were relatively similar (GM: −0.73% and NAWM: −0.79%; GM: $d=-0.46$ and NAWM: $d=-0.56$). Besides these small between-study differences, our results nevertheless concur with most cross-sectional and longitudinal studies.

Although the change of major brain metrics is similar to that reported in previous studies, our results for the cognitive performance measures we used here are different from several of the previous studies. We identified small but significant performance increases in the first 2 years, which were followed by a decline from the second to the fourth year. Although these changes were significant due to the relatively large sample size, they are particularly small, as indicated by the weak effect sizes (see Table 1). Thus, we could not replicate the cognitive decline in performance reported by other longitudinal studies (Ritchie et al 2015a, c; Persson et al 2016; Yuan et al 2018). We rather identified small performance increases across the first 2 years, followed by a small performance decrease in the fourth year. Interestingly, cognitive performance did not substantially decrease below the baseline level in the fourth year. In other words, cognitive performance is on average more or less stable in our cohort of elderly subjects. It could be argued that we did not detect cognitive performance declines due to the fact

that subjects become increasingly familiar with the test with each measurement. This issue would lead to procedural or even explicit learning effects. From our point of view, it is not entirely clear whether and how practice effects may have exerted an influence on the measured cognitive performance in our study due to several reasons. First, we must take into consideration that each examination was separated by at least 1 year and it is unlikely that the subjects remembered the details of the cognitive tests particularly well. Secondly, during each annual examination, the subjects were enrolled in several psychological and neuroimaging tests, making it unlikely that this particular test we are here relying upon is well remembered. Thirdly, although there is no doubt that practice effects are present in longitudinal cognitive studies, they mostly occur during the second measurement (accounting for 4–5% of the performance) while they are modest or even weak in subsequent measurements (Rabbitt et al 2004). Nevertheless, cognitive performance is more or less stable in our cohort of elderly subjects. In the level- and change–change analyses including the cognitive measures, we thus have modeled the assumed practice effects to detect level- and change–change associations more or less independent of practice effects (repeated exposure effects), similar to the approach adopted by Yuan et al. (2018).

Level–level and change–change associations for the brain metrics

As expected, we identified several level–level associations between the anatomical measures (see sTable 4 and Fig. 3). Although these associations are statistically significant due to the relatively large sample size, these associations are small according to Cohen's effect size classification. Only the level–level relation between GM and NAWM is large ($r = 0.56$) and in the range reported by previous studies ($r = 0.34$) (Cox et al. 2016). The other significant associations are much smaller (absolute r values ranging between 0.15 and 0.23) and will thus not be discussed intensively in the context of this paper. However, larger GM and NAWM volumes are related to smaller LV and WMH volumes likely indicating that larger GM and NAWM volumes prevent or mitigate the increase of LV and WMH volumes.

We also identified several significant and partly substantial level–change associations between the neuroanatomical measures. Subjects with large LV volumes at baseline demonstrated steep increases of LV and WMH volumes throughout the 4 years of examination. This result could indicate that those subjects already suffering from brain tissue loss at baseline (indicated by increased LV) may be those who will demonstrate a steeper progression of LV increase mostly likely related or even induced by brain tissue loss. On the other hand, large NAWM volumes at baseline are predictive for lesser increases of LV and WMH volumes. There was also a small but nevertheless significant positive level–change association between baseline NAWM and GM change. Larger NAWM volumes seem to be protective for brain tissue loss and even for WMH increases. The potential mechanisms explaining this relationship are currently unknown. However, it could be possible that subjects with larger brains and thus larger NAWM are slightly more intelligent and thus demonstrate better cognitive ability (Luders et al 2009), positively influencing their lifestyle. For example, it has been reported that subjects with better cognitive ability smoke less and exercise more (Gottfredson 2004), both of which may result in beneficial influences on brain structure, reduced WHMs, and other detrimental brain anomalies (Hillman et al 2008; Almeida et al 2011; Farooqui and Farooqui 2015). Another possible explanation is that the positive level–change associations between NAWM, GM, and WMH reflect a differential temporal sensitivity of these structures to a shared, underlying age-process—comparable to a positive manifold (Spearman 1904; Raz et al 2005). Future longer-term follow-ups may provide further insights into the sequential relationships between these brain structure changes.

There were also four significant change–change associations between the anatomical measures. WMH and LV volume increases are moderately and positively associated ($r = 0.47$). On the other hand, GM and NAWM changes over the course of 4 years are negatively associated with LV and WMH volume changes ($r = -0.53$ to $r = -0.6$). In other words, large volumes increases in WMH and LV are related to large decreases of GM and NAWM. These associations are not surprising since WMH and LV volume increases mostly happen at the expense of GM and NAWM volume decreases.

We identified several residual–residual associations between the trajectories of the anatomical measures additionally supporting a relatively strong interconnection between the trajectories of the anatomical measures. These were especially of a quite large magnitude for the associations between the residuals of LV to the residuals of the other brain metrics ($r = -0.47$ to GM, $r = 0.30$ to WMH, and $r = -0.25$ to NAWM). We do not expect that the brain changes are exactly the same per year for a person. The residual associations indicate that those specific yearly changes that are not captured by our model in one brain metric were related to the specific yearly changes of another brain metric. However, the residual-residual association might have been inflated and overestimated because all measures were obtained using the same FreeSurfer parcellation. This may explain the small but positive residual-residual association ($r = 0.15$) between the WMH and the NAWM.

Level–level, level–change, and change–change associations between brain metrics and cognitive performance

Relating the psychometric IQ measures to the brain metrics revealed mainly significant positive level–level associations between GM and IQ (language and non-language). NAWM was significantly and positively related to language processing. However, this relation was weak ($r = 0.18$) and became only significant due to the relatively large sample size. Weak (but negative) level–level associations were identified for LV. The positive level–level association of the IQ measures with GM ($r = 0.25$) is approximately similar to those reported in previous studies (Luders et al 2009); however, the association we found in our study is at the lower edge of reported associations, with the majority of studies reporting correlations around $r = 0.3$–0.4 (cited from Luders et al 2009). Thus, larger GM volume is weakly related to higher IQ. However, the small level–level relations for NAWM and LV as well as the non-existent associations between cognitive performance and WMH do not replicate the findings of several previous studies. Most of these studies report correlations between WM and IQ or other cognitive functions in the same range as for GM. Although the small correlation between GM and IQ again supports the idea that GM (and thus supposedly the larger number of neurons as well as the larger neuropil) may be related to human intelligence and cognition, other factors must also play a role. For example, functional connectivity patterns (Langer et al 2012), the efficiency of neurophysiological activation (Neubauer and Fink 2009), or particular neurophysiological oscillation patterns (Thatcher et al 2008) may account for even larger portions of variance when assessing relationships between brain measures and intelligence.

There were only weak and non-significant level–change associations between language IQ and LV as well as WMH. The higher the language IQ, the weaker the increases in LV and WMH over the 4 years. Most important, we did not find evidence for change–change associations between the IQ and the anatomical brain measures. In this respect, our findings also contribute to the inhomogeneous results of earlier studies. As yearly changes in intelligence are small compared to the measurement error of the IQ tests, high uncertainty is involved in associating the brain metrics to the person-specific slopes of the IQ measures. The uncertainty was highest associating the slopes of GM and NAWM to the slopes of the IQ measures because the person-specific slopes of GM and NAWM also did not vary substantially.

Of the few longitudinal studies investigating associations of GM with IQ in healthy older adults, four reported significant level–change or correlated change relationships between these measures (Raz et al 2007, 2008; Ritchie et al 2015a; Persson et al 2016; Yuan et al 2018), while two did not find any effect (Raz et al 2007; Gorbach et al 2017). Regarding level–change associations, Persson et al. (2016) found that higher GM volume at baseline was related to reduced 2-year changes in fluid IQ. Moreover, two other studies observed that lower baseline fluid IQ was related to lower GM volume decline (Raz et al 2008; Yuan et al 2018). Interestingly, Yuan, et al. (2018) additionally found a negative level–change association between crystallized IQ and cortical GM volume decline, suggesting that fluid and crystallized IQ are differentially related to age-related GM atrophy. Only one study reported significant positive correlated change associations between total GM volume and IQ (Ritchie et al 2015c). Of note, this study is also the only one that included exclusively older adults within a particularly narrow age range (i.e., participants aged 73). While these predominantly positive level–change and correlated change relationships suggest a protective effect of preserved GM volumes or IQ for age-related variations in the respective other domain, it is difficult to disentangle the directionality of these associations.

Second, several previous studies investigated level–change or correlated change associations between WMH or NAWM and a measure of IQ (Schmidt et al 1999, 2005; Raz et al 2007; Ritchie et al 2015a, 2015c; Persson et al 2016; Gorbach et al 2017). Of these, three studies reported negative correlated changes between WMH and fluid IQ (Schmidt et al 2005; Raz et al 2007; Ritchie et al 2015a), suggesting that an increase in WMH is related to a simultaneous IQ decline, while three studies did not find any level–change or correlated change relationships (Schmidt et al 1999; Raz et al 2008; Gorbach et al 2017). With regard to NAWM, Persson et al. (2016) found a positive association between lower baseline prefrontal WM volume and prefrontal WM atrophy with subsequent 2-year degradation in IQ in a sample of healthy younger and older adults (aged between 19 and 79). However, two other studies did not report any significant level–change or correlated change relationship of NAWM and IQ (Raz et al 2008; Ritchie et al 2015a).

Finally, to the best of our knowledge, no previous longitudinal study has directly examined longitudinal associations between LV and IQ in healthy older adults. In a sample of middle-aged to older adults, Leong et al. (2017) have examined 8-year changes in total ventricle volume and in a composite of global cognition (average of EF, processing speed, memory and attention tasks) that might be comparable to an estimate of IQ. However, the authors did not find a relationship between change in total ventricle volume and change in global cognition.

Overall, the heterogeneity of previous findings between age-related structural brain and IQ changes is most likely in part due to the broad age ranges used in the majority of these studies, the inclusion of different measures of intelligence, and the wide variety of statistical methods employed.

That we did not find any reliable level–change or correlated change relationships between verbal/nonverbal IQ and a number of different brain measures may also be a reflection of the exceptionally high health status of our participants, who have an above-average level of education compared to the Swiss population (Bundesamt für Statistik 2017) and are cognitively and physically fit. According to well-established cognitive aging theories (Stern 2002, 2009; Park and Reuter-Lorenz 2009; Reuter-Lorenz and Park 2014), healthy older adults are assumed to have sufficient resources to compensate for age-related brain degradation for a certain time and thus delay or at least attenuate a decrease in IQ. It might, therefore, be more likely that change–change relationships between brain structure and cognitive abilities follow a lagged pattern. For example, in another analysis of this sample, we did not find any evidence for simultaneous correlated changes between WM microstructure and processing speed over 4 years; however, there was some evidence for a lagged change relationship, such that changes in WM microstructure preceded changes in processing speed (Oschwald et al 2019b).

Limitations

In the present study, we used data from a comparatively large sample of healthy older adults ($N = 231$) given the standards of other longitudinal studies investigating level–change or correlated change associations between structural brain and cognitive measures. Furthermore, we included four measurement occasions, an approach that allowed us to model change trajectories more precisely, also testing for quadratic changes over time. However, in light of the complexity of multivariate growth models such as the ones used in the present analysis, an even larger sample size and more measurement occasions would ideally be available. For example, we did not have a sufficiently large number of observations to model level- and correlated–change relationships between all brain measures and the two IQ measures simultaneously in one multivariate model, and we, therefore, instead investigated change-relationships of GM/NAWM and WMH/LV separately with the IQ measures. With longitudinal neuroimaging studies such as the present one already including (e.g., Seattle Longitudinal Study Schaie 1996; Schaie and Willis 2010) or planning to include more than two measurement occasions in the future (e.g., Harvard Aging Brain Study or Cambridge Center for Ageing Neuroscience Study Shafto et al 2014; Dagley et al 2017), the use of data pooling across multiple sites provides a promising avenue to increase sample size and gain insights into the generalizability of level–change and correlated–change relationships between brain and behavior (Jockwitz et al 2019).

While we included a number of different brain metrics to provide multifaceted insights into the interplay of different brain tissue changes in old age and their relation with IQ changes, these measures were all relatively broad. Previous research has suggested that IQ relies on distributed networks in the brain (Jung and Haier 2007; Langer et al 2012). It is, therefore, possible that the inclusion of more specific brain metrics from these networks (regional structural brain metrics, indices of network connectivity) may have yielded more reliable level–change and correlated change relationships with IQ.

With regard to the brain metrics used, another limitation that requires mentioning is the method we used for the automated parcellation of the brain volumes and WMH. While the FreeSurfer software generally yields reliable measurements across retest intervals (Liem et al 2015), the retest-reliability is dependent on the type and size of the reconstructed brain structure (Morey et al 2010). Furthermore, the accuracy of WMH segmentation using automated methods is currently unclear because the distribution of WMHs in the brain is insufficiently understood (Caligiuri et al 2015) and the standard approach for segmenting WM lesions usually relies on FLAIR imaging data, not used in the present study.

An advantage of the present study is that we included a sample of exclusively older adults (above age 64 years), and thus were better able to capture age-related changes than in previous studies covering the entire adult lifespan. Yet, our age range was still relatively wide (64–86 years at baseline) compared to studies limiting their investigation to similar-aged adults (Ritchie et al 2015a, c). It is highly likely that different aging processes on the level of the brain and cognitive ability occur in an individual of 64 years compared to somebody of more advanced age. Thus, besides age-related within-person changes that were of main interest in this study, our findings might also be influenced by a certain amount of between-person age-differences.

We are aware of the fact that several previous studies based their analysis on the IQ model originally proposed by Charles Spearman and Raymond Cattell postulating that intelligence would not be a unitary entity. They introduced the concepts of fluid (Gf) and crystallized intelligence (Gc) as independent components (Cattell 1943, 1963). Gf refers to the capacity for logical reasoning and problem-solving that is presumably independent of acquired knowledge while Gc stands for the ability to use acquired and culture-relevant information. Although Gf and Gc are often viewed as distinct intelligence components, several studies and theoretical contentions suggest that they are not statistically independent (Carroll 1993). For example, Thurstone (1938) proposed, based on his own empirical analyses, seven primary mental abilities and not only two independent intelligence factors. In our project, we did not use the Gf and Gc concepts but rather worked with psychometric IQ measures based on the Thurston model because the IQ test we have used has been specifically adapted to examine subjects at the age of 50–90 years.

Conclusion

Examining a large longitudinal sample of cognitively healthy older adults across 4 years, we found changes in several brain anatomical measures (GM, NAWM, WMH, LV) largely comparable to the findings of previous cross-sectional and longitudinal studies. In contrast to recent evidence, however, change in verbal and nonverbal IQ measures in our sample remained relatively stable. In addition, we found evidence for level–level and -change associations between several of the brain anatomical measures. With regard to brain-cognition associations, the only substantial finding was a positive level–level association between GM, NAWM, and the IQ measures, indicating that participants with larger brain volumes had a higher IQ. However, no (or only weak) evidence was found for level- or change–change associations between any of the brain metrics and the IQ measures. Taken together, these results suggest that while healthy older adults show an age-related neuroanatomical decline over a time span of 4 years, these degenerative changes are not necessarily linked to simultaneous cognitive deterioration.

Acknowledgements The current analysis incorporates data from the Longitudinal Healthy Aging Brain (LHAB) database project, which is carried out as one of the core projects at the International Normal Aging and Plasticity Imaging Center/INAPIC and the University Research Priority Program "Dynamics of Healthy Aging" of the University of Zurich. The following members of the core INAPIC team were involved in the design, set-up, maintenance, and support of the LHAB database: Anne Eschen, Lutz Jäncke, Mike Martin, Susan Merillat, Christina Rocke, and Jacqueline Zöllig.

References

Almeida OP, Garrido GJ, Alfonso H et al (2011) 24-month effect of smoking cessation on cognitive function and brain structure in later life. Neuroimage 55:1480–1489. https://doi.org/10.1016/j.neuroimage.2011.01.063

Anblagan D, Valdés Hernández MC, Ritchie SJ et al (2018) Coupled changes in hippocampal structure and cognitive ability in later life. Brain Behav 8:e00838. https://doi.org/10.1002/brb3.838

Bangen KJ, Werhane ML, Weigand AJ et al (2018) Reduced regional cerebral blood flow relates to poorer cognition in older adults with type 2 diabetes. Front Aging Neurosci 10:270. https://doi.org/10.3389/fnagi.2018.00270

Baraldi AN, Enders CK (2010) An introduction to modern missing data analyses. J Sch Psychol 48:5–37. https://doi.org/10.1016/j.jsp.2009.10.001

Bartlett JW, Frost C (2008) Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound Obstet Gynecol 31:466–475. https://doi.org/10.1002/uog.5256

Bender AR, Völkle MC, Raz N (2016) Differential aging of cerebral white matter in middle-aged and older adults: a seven-year follow-up. Neuroimage 125:74–83. https://doi.org/10.1016/j.neuroimage.2015.10.030

Bland JM, Altman DG (2003) Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol 22:85–93. https://doi.org/10.1002/uog.122

Bundesamt für Statistik (2017) Bildungsstand der ständigen Wohnbevölkerung nach Alter und Geschlecht. https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/tabellen.assetdetail.7886041.html

Caligiuri ME, Perrotta P, Augimeri A et al (2015) Automatic detection of white matter hyperintensities in healthy aging and pathology using magnetic resonance imaging: a review. Neuroinformatics 13:261–276. https://doi.org/10.1007/s12021-015-9260-y

Carroll J (1993) Human cognitive abilities: a survey of factor-analytic studies. Cambridge University Press, Cambridge

Catchlove SJ, Macpherson H, Hughes ME et al (2018) An investigation of cerebral oxygen utilization, blood flow and cognition in healthy aging. PLoS ONE 13:e0197055. https://doi.org/10.1371/journal.pone.0197055

Cattell RB (1943) The measurement of adult intelligence. Psychol Bull 40:153–193. https://doi.org/10.1037/h0059973

Cattell RB (1963) Theory of fluid and crystallized intelligence: a critical experiment. J Educ Psychol 54:1–22. https://doi.org/10.1037/h0046743

Charlton RA, Schiavone F, Barrick TR et al (2010) Diffusion tensor imaging detects age related white matter change over a 2 year follow-up which is associated with working memory decline. J Neurol Neurosurg Psychiatry 81:13–19

Cohen RM, Small C, Lalonde F et al (2001) Effect of apolipoprotein E genotype on hippocampal volume loss in aging healthy women. Neurology 57:2223–2228. https://doi.org/10.1212/wnl.57.12.2223

Cox SR, Ritchie SJ, Tucker-Drob EM, Liewald DC, Hagenaars SP, Davies G, Wardlaw JM, Gale CR, Bastin Deary IJ (2016) Ageing and brain white matter structure in 3,513 UK Biobank participants. Nature Communications 7:13629. https://doi.org/10.1038/ncomms13629

Cribbie RA (2007) Multiplicity control in structural equation modeling. Struct Equ Model 14:98–112

Dagley A, LaPoint M, Huijbers W et al (2017) Harvard aging brain study: dataset and accessibility. Neuroimage 144:255–258. https://doi.org/10.1016/j.neuroimage.2015.03.069

Daugherty AM, Haacke EM, Raz N (2015) Striatal iron content predicts its shrinkage and changes in verbal working memory after two years in healthy adults. J Neurosci 35:6731–6743. https://doi.org/10.1523/JNEUROSCI.4717-14.2015

Destrieux C, Fischl B, Dale A, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53:1–15. https://doi.org/10.1016/j.neuroimage.2010.06.010

Farooqui AA, Farooqui T (2015) Diet and Exercise in Cognitive Function and Neurological Diseases. John Wiley & sons

Fischl B, Salat DH, Busa E et al (2002) Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33:341–355. https://doi.org/10.1016/s0896-6273(02)00569-x

Fischl B, van der Kouwe A, Destrieux C et al (2004) Automatically parcellating the human cerebral cortex. Cereb Cortex 14:11–22. https://doi.org/10.1093/cercor/bhg087

Fjell AM, Walhovd KB (2010) Structural brain changes in aging: courses, causes and cognitive consequences. Rev Neurosci 21:187–221

Fjell AM, Westlye LT, Grydeland H et al (2014) Accelerating cortical thinning: unique to dementia or universal in aging? Cereb Cortex 24:919–934. https://doi.org/10.1093/cercor/bhs379

Fjell AM, Sneve MH, Storsve AB et al (2016) Brain events underlying episodic memory changes in aging: a longitudinal investigation of structural and functional connectivity. Cereb Cortex 26:1272–1286. https://doi.org/10.1093/cercor/bhv102

Fjell AM, Sneve MH, Grydeland H et al (2017) Relationship between structural and functional connectivity change across the adult lifespan: a longitudinal investigation. Hum Brain Mapp 38:561–573. https://doi.org/10.1002/hbm.23403

Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Gorbach T, Pudas S, Lundquist A et al (2017) Longitudinal association between hippocampus atrophy and episodic-memory decline. Neurobiol Aging 51:167–176. https://doi.org/10.1016/j.neurobiolaging.2016.12.002

Gorgolewski KJ, Alfaro-Almagro F, Auer T et al (2017) BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. PLoS Comput Biol 13:e1005209. https://doi.org/10.1371/journal.pcbi.1005209

Gottfredson LS (2004) Intelligence: is it the epidemiologists' elusive "fundamental cause" of social class inequalities in health? J Pers Soc Psychol 86:174–199. https://doi.org/10.1037/0022-3514.86.1.174

Grimm KJ, An Y, McArdle JJ et al (2012) Recent changes leading to subsequent changes: extensions of multivariate latent difference score models. Struct Equ Model 19:268–292. https://doi.org/10.1080/10705511.2012.659627

Grimm KJ, Ram N, Estabrook R (2016) Growth modeling: structural equation and multilevel modeling approaches. Guilford Publications, New York

Hillman CH, Erickson KI, Kramer AF (2008) Be smart, exercise your heart: exercise effects on brain and cognition. Nat Rev Neurosci 9:58–65. https://doi.org/10.1038/nrn2298

Hirsiger S, Koppelmans V, Mérillat S et al (2016) Structural and functional connectivity in healthy aging: associations for cognition and motor behavior. Hum Brain Mapp 37:855–867. https://doi.org/10.1002/hbm.23067

Hohman TJ, Tommet D, Marks S et al (2017) Evaluating Alzheimer's disease biomarkers as mediators of age-related cognitive decline. Neurobiol Aging 58:120–128. https://doi.org/10.1016/j.neurobiolaging.2017.06.022

Jäncke L, Mérillat S, Liem F, Hänggi J (2015) Brain size, sex, and the aging brain. Hum Brain Mapp 36:150–169. https://doi.org/10.1002/hbm.22619

Jockwitz C, Mérillat S, Liem F et al (2019) Generalizing age effects on brain structure and cognition: a two-study comparison approach. Hum Brain Mapp 40:2305–2319. https://doi.org/10.1002/hbm.24524

Jung RE, Haier RJ (2007) The Parieto-Frontal integration theory (P-FIT) of intelligence: converging neuroimaging evidence. Behav Brain Sci 30:135–154. https://doi.org/10.1017/S0140525X07001185**(discussion 154)**

King KM, Littlefield AK, McCabe CJ et al (2018) Longitudinal modeling in developmental neuroimaging research: common challenges, and solutions from developmental psychology. Dev Cogn Neurosci 33:54–72. https://doi.org/10.1016/j.dcn.2017.11.009

Köhncke Y, Laukka EJ, Brehmer Y et al (2016) Three-year changes in leisure activities are associated with concurrent changes in white matter microstructure and perceptual speed in individuals aged 80 years and older. Neurobiol Aging 41:173–186. https://doi.org/10.1016/j.neurobiolaging.2016.02.013

Langer N, Pedroni A, Gianotti LRR et al (2012) Functional brain network efficiency predicts intelligence. Hum Brain Mapp 33:1393–1406. https://doi.org/10.1002/hbm.21297

Leong RLF, Lo JC, Sim SKY et al (2017) Longitudinal brain structure and cognitive changes over 8 years in an East Asian cohort. Neuroimage 147:852–860. https://doi.org/10.1016/j.neuroimage.2016.10.016

Leow AD, Yanovsky I, Parikshak N et al (2009) Alzheimer's disease neuroimaging initiative: a one-year follow up study using tensor-based morphometry correlating degenerative rates, biomarkers and cognition. Neuroimage 45:645–655. https://doi.org/10.1016/j.neuroimage.2009.01.004

Liem F, Mérillat S, Bezzola L et al (2015) Reliability and statistical power analysis of cortical and subcortical FreeSurfer metrics in a large sample of healthy elderly. Neuroimage 108:95–109. https://doi.org/10.1016/j.neuroimage.2014.12.035

Lövdén M, Köhncke Y, Laukka EJ et al (2014) Changes in perceptual speed and white matter microstructure in the corticospinal tract are associated in very old age. Neuroimage 102(Pt 2):520–530. https://doi.org/10.1016/j.neuroimage.2014.08.020

Luders E, Narr KL, Bilder RM et al (2007) Positive correlations between corpus callosum thickness and intelligence. Neuroimage 37:1457–1464. https://doi.org/10.1016/j.neuroimage.2007.06.028

Luders E, Narr KL, Bilder RM et al (2008) Mapping the relationship between cortical convolution and intelligence: effects of gender. Cereb Cortex 18:2019–2026. https://doi.org/10.1093/cercor/bhm227

Luders E, Narr KL, Thompson PM, Toga AW (2009) Neuroanatomical correlates of intelligence. Intelligence 37:156–163. https://doi.org/10.1016/j.intell.2008.07.002

Madhyastha T, Mérillat S, Hirsiger S et al (2014) Longitudinal reliability of tract-based spatial statistics in diffusion tensor imaging. Hum Brain Mapp 35:4544–4555. https://doi.org/10.1002/hbm.22493

Mak E, Su L, Williams GB et al (2015a) Longitudinal assessment of global and regional atrophy rates in Alzheimer's disease and dementia with Lewy bodies. Neuroimage Clin 7:456–462. https://doi.org/10.1016/j.nicl.2015.01.017

Mak E, Su L, Williams GB et al (2015b) Progressive cortical thinning and subcortical atrophy in dementia with Lewy bodies and Alzheimer's disease. Neurobiol Aging 36:1743–1750. https://doi.org/10.1016/j.neurobiolaging.2014.12.038

Mathalon DH, Sullivan EV, Rawles JM, Pfefferbaum A (1993) Correction for head size in brain-imaging measurements. Psychiatry Res Neuroimaging 50:121–139. https://doi.org/10.1016/0925-4927(93)90016-B

McArdle JJ, Hamgami F, Jones K et al (2004) Structural modeling of dynamic changes in memory and brain structure using longitudinal data from the normative aging study. J Gerontol B Psychol Sci Soc Sci 59:P294–304. https://doi.org/10.1093/geronb/59.6.p294

Miraglia F, Vecchio F, Bramanti P, Rossini PM (2016) EEG characteristics in "eyes-open" versus "eyes-closed" conditions: small-world network architecture in healthy aging and age-related brain degeneration. Clin Neurophysiol 127:1261–1268. https://doi.org/10.1016/j.clinph.2015.07.040

Möller C, Hafkemeijer A, Pijnenburg YAL et al (2016) Different patterns of cortical gray matter loss over time in behavioral variant frontotemporal dementia and Alzheimer's disease. Neurobiol Aging 38:21–31. https://doi.org/10.1016/j.neurobiolaging.2015.10.020

Moon SY, de Souto BP, Chupin M et al (2017) Associations between white matter hyperintensities and cognitive decline over three years in non-dementia older adults with memory complaints. J Neurol Sci 379:266–270. https://doi.org/10.1016/j.jns.2017.06.031

Morey RA, Selgrade ES, Wagner HR et al (2010) Scan-rescan reliability of subcortical brain volumes derived from automated segmentation. Hum Brain Mapp 31:1751–1762. https://doi.org/10.1002/hbm.20973

Neubauer AC, Fink A (2009) Intelligence and neural efficiency. Neurosci Biobehav Rev 33:1004–1023. https://doi.org/10.1016/j.neubiorev.2009.04.001

Oschwald J, Guye S, Liem F et al (2019a) Brain structure and cognitive ability in healthy aging: a review on longitudinal correlated change. Rev Neurosci. https://doi.org/10.1515/revneuro-2018-0096

Oschwald J, Mérillat S, Liem F, Röcke C, Martin M, Jäncke L (2019b) Lagged coupled changes between white matter microstructure and processing speed in healthy aging: a longitudinal investigation. Front Aging Neurosci 11:1–19. https://doi.org/10.3389/fnagi.2019.00298

Park DC, Reuter-Lorenz P (2009) The adaptive brain: aging and neurocognitive scaffolding. Annu Rev Psychol 60:173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656

Persson J, Pudas S, Lind J et al (2012) Longitudinal structure-function correlates in elderly reveal MTL dysfunction with cognitive decline. Cereb Cortex 22:2297–2304. https://doi.org/10.1093/cercor/bhr306

Persson N, Ghisletta P, Dahle CL et al (2016) Regional brain shrinkage and change in cognitive performance over two years: the bidirectional influences of the brain and cognitive reserve factors. Neuroimage 126:15–26. https://doi.org/10.1016/j.neuroimage.2015.11.028

Rabbitt P, Diggle P, Holland F, McInnes L (2004) Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. J Gerontol B Psychol Sci Soc Sci 59:P84–97

Raz N, Lindenberger U, Rodrigue KM et al (2005) Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. Cereb Cortex 15:1676–1689. https://doi.org/10.1093/cercor/bhi044

Raz N, Rodrigue KM, Kennedy KM, Acker JD (2007) Vascular health and longitudinal changes in brain and cognition in middle-aged and older adults. Neuropsychology 21:149–157. https://doi.org/10.1037/0894-4105.21.2.149

Raz N, Lindenberger U, Ghisletta P et al (2008) Neuroanatomical correlates of fluid intelligence in healthy adults and persons with vascular risk factors. Cereb Cortex 18:718–726. https://doi.org/10.1093/cercor/bhm108

Reuter-Lorenz PA, Park DC (2014) How does it STAC up? Revisiting the scaffolding theory of aging and cognition. Neuropsychol Rev 24:355–370. https://doi.org/10.1007/s11065-014-9270-9

Ritchie SJ, Bastin ME, Tucker-Drob EM et al (2015a) Coupled changes in brain white matter microstructure and fluid intelligence in later life. J Neurosci 35:8672–8682. https://doi.org/10.1523/JNEUROSCI.0862-15.2015

Ritchie SJ, Booth T, Valdés Hernández MDC et al (2015b) Beyond a bigger brain: multivariable structural brain imaging and intelligence. Intelligence 51:47–56. https://doi.org/10.1016/j.intell.2015.05.001

Ritchie SJ, Dickie DA, Cox SR et al (2015c) Brain volumetric changes and cognitive ageing during the eighth decade of life. Hum Brain Mapp 36:4910–4925. https://doi.org/10.1002/hbm.22959

Rosseel Y (2012) lavaan: an R Package for structural equation modeling. J Stat Softw. https://doi.org/10.18637/jss.v048.i02

Sala-Llonch R, Idland A-V, Borza T et al (2017) Inflammation, amyloid, and atrophy in the aging brain: relationships with longitudinal changes in cognition. J Alzheimers Dis 58:829–840. https://doi.org/10.3233/JAD-161146

Salthouse TA (2010) Selective review of cognitive aging. J Int Neuropsychol Soc 16:754–760. https://doi.org/10.1017/S1355617710000706

Schaie KW (1996) Intellectual development in adulthood: the Seattle longitudinal study. Cambridge University Press, Cambridge

Schaie KW, Willis SL (2010) The seattle longitudinal study of adult cognitive development. ISSBD Bull 57:24–29

Schmidt R, Fazekas F, Kapeller P et al (1999) MRI white matter hyperintensities: three-year follow-up of the Austrian Stroke Prevention Study. Neurology 53:132–139. https://doi.org/10.1212/wnl.53.1.132

Schmidt R, Ropele S, Enzinger C et al (2005) White matter lesion progression, brain atrophy, and cognitive decline: the Austrian stroke prevention study. Ann Neurol 58:610–616. https://doi.org/10.1002/ana.20630

Shafto MA, Tyler LK, Dixon M et al (2014) The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. BMC Neurol 14:204. https://doi.org/10.1186/s12883-014-0204-1

Silbert LC, Nelson C, Howieson DB et al (2008) Impact of white matter hyperintensity volume progression on rate of cognitive and motor decline. Neurology 71:108–113. https://doi.org/10.1212/01.wnl.0000316799.86917.37

Sleimen-Malkoun R, Perdikis D, Müller V et al (2015) Brain Dynamics of Aging: Multiscale Variability of EEG Signals at Rest and during an Auditory Oddball Task. eNeuro. https://doi.org/10.1523/ENEURO.0067-14.2015

Smith CE, Cribbie RA (2013) Multiplicity control in structural equation modeling: incorporating parameter dependencies. Struct Equat Model 20:79–85. https://doi.org/10.1080/10705511.2013.742385

Song Z, Farrell ME, Chen X, Park DC (2018) Longitudinal accrual of neocortical amyloid burden is associated with microstructural changes of the fornix in cognitively normal adults. Neurobiol Aging 68:114–122. https://doi.org/10.1016/j.neurobiolaging.2018.02.021

Spearman C (1904) "General Intelligence", objectively determined and measured. Am J Psychol 15:201. https://doi.org/10.2307/1412107

Stern Y (2002) What is cognitive reserve? Theory and research application of the reserve concept. J Int Neuropsychol Soc 8:448–460. https://doi.org/10.1017/S1355617702813248

Stern Y (2009) Cognitive reserve. Neuropsychologia 47:2015–2028. https://doi.org/10.1016/j.neuropsychologia.2009.03.004

Sturm W, Horn W, Willmes K (1993) Leistungsprüfsystem für 50–90jährige:(LPS 50+). Hogrefe, Verlag für Psychologie, Handanweisung

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Thatcher RW, North DM, Biver CJ (2008) Intelligence and EEG phase reset: a two compartmental model of phase shift and lock. Neuroimage 42:1639–1653. https://doi.org/10.1016/j.neuroimage.2008.06.009

Thurstone LL (1938) Primary mental abilities. University of Chicago, Chicago, p 121

Valizadeh SA, Hänggi J, Mérillat S, Jäncke L (2017) Age prediction on the basis of brain anatomical measures. Hum Brain Mapp 38:997–1008. https://doi.org/10.1002/hbm.23434

Valizadeh SA, Liem F, Mérillat S et al (2018) Identification of individual subjects on the basis of their brain anatomical features. Sci Rep 8:5611. https://doi.org/10.1038/s41598-018-23696-6

Voyer D (1996) On the magnitude of laterality effects and sex differences in functional lateralities. Laterality 1:51–83

Walhovd KB, Westlye LT, Amlien I et al (2011) Consistent neuroanatomical age-related volume differences across multiple samples. Neurobiol Aging 32:916–932. https://doi.org/10.1016/j.neurobiolaging.2009.05.013

Ware JE, Kosinski M, Keller SD (1996) A 12-item short-form health survey. Med Care 34:220–233. https://doi.org/10.1097/00005650-199603000-00003

Ware JE, Kosinski M, Keller SD, Lincoln RI (1996b) Sf-12: How to score the sf-12 physical and mental health summary scales. Health Institute, Boston

Yuan P, Voelkle MC, Raz N (2018) Fluid intelligence and gross structural properties of the cerebral cortex in middle-aged and older adults: a multi-occasion longitudinal study. Neuroimage 172:21–30. https://doi.org/10.1016/j.neuroimage.2018.01.032

Zöllig J, Mérillat S, Eschen A et al (2011) Plasticity and imaging research in healthy aging: core ideas and profile of the International Normal Aging and Plasticity Imaging Center (INAPIC). Gerontology 57:190–192. https://doi.org/10.1159/000324307