Cerebral Cortex, 2020;00: 1–11

doi: 10.1093/cercor/bhaa296
Original Article

ORIGINAL ARTICLE

The Speed of Development of Adolescent Brain Age Depends on Sex and Is Genetically Determined

Rachel M. Brouwer[1], Jelle Schutte[1], Ronald Janssen[1], Dorret I. Boomsma[2], Hilleke E. Hulshoff Pol[1] and Hugo G. Schnack[1]

[1]Department of Psychiatry, University Medical Center Utrecht Brain Center, Utrecht University, 3584 CX Utrecht, the Netherlands and [2]Department of Biological Psychology and Netherlands Twin Register, VU University Amsterdam, 1081 HV Amsterdam, the Netherlands

Address correspondence to Dr Rachel Brouwer. Email: r.m.brouwer-4@umcutrecht.nl.

Abstract

Children and adolescents show high variability in brain development. Brain age—the estimated biological age of an individual brain—can be used to index developmental stage. In a longitudinal sample of adolescents (age 9–23 years), including monozygotic and dizygotic twins and their siblings, structural magnetic resonance imaging scans ($N = 673$) at 3 time points were acquired. Using brain morphology data of different types and at different spatial scales, brain age predictors were trained and validated. Differences in brain age between males and females were assessed and the heritability of individual variation in brain age gaps was calculated. On average, females were ahead of males by at most 1 year, but similar aging patterns were found for both sexes. The difference between brain age and chronological age was heritable, as was the change in brain age gap over time. In conclusion, females and males show similar developmental ("aging") patterns but, on average, females pass through this development earlier. Reliable brain age predictors may be used to detect (extreme) deviations in developmental state of the brain early, possibly indicating aberrant development as a sign of risk of neurodevelopmental disorders.

Key words: brain age, heritability, longitudinal imaging, sex differences, structural brain development

Introduction

Brain morphology is continuously changing throughout life. One way of quantifying individual differences in the context of development and aging is to compute a "brain age," which is a neurobiological marker of age, rather than a chronological one. Brain age is a predicted age based on brain phenotypes such as brain structure (structural magnetic resonance imaging [MRI]; Franke et al. 2010) or functioning (functional MRI; Dosenbach et al. 2010), usually determined through machine learning techniques. Differences between predicted and chronological age can be interpreted as delayed or accelerated development or aging ("brain age gap"). Advanced brain age has been found for several brain diseases, such as schizophrenia (Hajek et al. 2019; Koutsouleris et al. 2014; Nenadić et al. 2017; Schnack et al. 2016), and Alzheimer's disease (Gaser et al. 2013; Lowe et al. 2016). Since psychotic disorders originate in adolescence, advanced brain age could be used as a clinically useful marker for early aberrant development. For example, increased brain age has been found in subjects at high risk for psychosis (Koutsouleris et al. 2014). Since structural brain changes in childhood and adolescence are expected to differ qualitatively and quantitatively from those in adulthood, a brain age model tailored to the younger age range would likely be more accurate and more sensitive to subtle aging effects in adolescents at risk. Several studies have been carried

out in children and adolescents thus far, showing a mean absolute prediction error of around 1.0–1.7 years (Brown et al. 2012; Cao et al. 2015; Erus et al. 2015; Franke et al. 2012; Khundrakpam et al. 2015; Truelove-Hill et al. 2020). These deviations from the chronological age can be interpreted as delayed or accelerated brain development, making brain age a potential biomarker to detect deviant development.

The high heritability of both brain structure and psychiatric disorders gives rise to the question whether genes drive individual differences in brain age gaps, as this would further validate the use of brain age as a risk factor for disease. Brain age itself has been found to be a heritable phenotype in a small sample of adult females (Cole et al. 2017). The question is whether delayed or accelerated brain age during development can be explained by genetic factors. Heritability of adolescent brain age gaps from functional data has recently been established based on electroencephalography (EEG)-derived brain measures (Vandenbosch et al. 2019). In that study, the EEG-based brain age gap was relatively stable over a period from childhood to adolescence. The MRI-based maturation index, a measure related to brain age, has also been found to be relatively stable throughout development (Cao et al. 2015). It remains an open question whether the temporal dynamics of brain age gaps over time are genetically driven.

In this work, we study several factors that could explain individual differences in brain age gaps. Sex may be one of those factors: females have been found to be more advanced in their development of brain structure (e.g., Gennatas et al. 2017; Giedd et al. 1999; Koolschijn and Crone 2013). In addition, females reach puberty earlier than males by a few years (Mul et al. 2001), and pubertal hormone levels have been associated with brain structure (reviewed in Herting and Sowell, 2017; Vijayakumar et al. 2018). We hypothesize that an MRI-based age prediction model in adolescence will estimate the brains of females to be older than their chronological age and those of males to be younger than their chronological age. Second, we investigate the influence of genetic factors on brain age gaps derived from MRI and on the dynamics of this brain age gap over time. Finally, we investigate brain age gaps derived from different neurobiological measures at different scales to investigate whether these capture the same underlying developmental processes. We base the age prediction on three different feature sets: volumetric region-of-interest (ROI) measures, vertex-based cortical thickness measures, and voxel-based gray matter density (GMD). We study these questions in a longitudinal, extended twin cohort that was followed throughout puberty and adolescence (Supplementary Table S1, Koenis et al. 2018; van Soelen et al. 2012a; Peper et al. 2009). We build sex-specific brain age models to investigate the influence of sex. We apply longitudinal twin modeling to disentangle the influences of genetic and environmental factors that act on the brain age gaps and to assess heritability of the dynamics of the brain age gaps over time. Finally, we investigate the (genetic) overlap between MRI-based brain age gaps derived at different spatial scales.

Methods and Materials

Participants

A total of 330 subjects from 112 families consisting of twin pairs and their siblings participated in the longitudinal BrainSCALE study on brain and cognitive development during childhood and adolescence (van Soelen et al. 2012b), a collaborative project between the Netherland Twin Register (Boomsma et al. 2006; Van Beijsterveldt et al. 2013) at the Vrije Universiteit Amsterdam and the University Medical Center Utrecht (UMCU). The BrainSCALE cohort consists of healthy typically developing children. Exclusion criteria for participation at the start of the study were: a known major medical or psychiatric history; chronic use of medication; participation in special education; physical or sensory disabilities; having a pacemaker; metal materials in the head with the exception for dental braces (van Soelen et al. 2012b). At follow-up, children were not assessed for these exclusion criteria (except for contraindications for MRI). The families first participated in 2005 when the twins were 9.1 (0.1) years of age and revisited the UMCU when the twins were 12.1 (0.2) and 17.2 (0.2) years old. The siblings were 11.8 (1.1), 14.8 (1.3), and 19.8 (1.3) years old at baseline and follow-ups, respectively. The total age range of the sample was 9.0 to 22.9 years. Return rates were 80% and 77% of initial inclusion. All members of a family were scanned on the same day (96%) with the exception of 11 occasions (2 twins, 9 siblings). In these cases, the participant was scanned within 2 months of their siblings. See Supplementary Table S1 for demographics of participants included in this analysis. This study was approved by the Central Committee on Research Involving Human Subjects of Netherlands (CCMO), and studies were performed in accordance with the Declaration of Helsinki. Parents signed informed consent forms for their children. At the third measurement, the adolescents signed their own informed consent forms. Parents were financially compensated for travel expenses, and children received a present or gift voucher at the end of the testing days.

MRI Acquisition

Participants underwent MRI on two identical 1.5 Tesla Philips Achieva scanners (Philips) at the UMCU. For measures of brain morphology, a three-dimensional T_1-weighted scan (spoiled gradient echo; time echo = 4.6 ms; time repetition = 30 ms; flip angle = 30°; 160–180 contiguous coronal slices of 1.2 mm; in-plane resolution of 1.0×1.0 mm^2; acquisition matrix of 256×256 voxels; field-of-view of 256 mm with 70% scan percentage) of the whole head was acquired (Peper et al. 2009) with the same acquisition parameters and on the same scanners for each visit. There was considerable dropout at MRI scanning at the second visit (24%), mainly due to dental braces. In addition, a number of scans dropped out because of motion (11%/10%/7%). In total, this study contained 673 scans from 305 subjects (Supplementary Table S1).

Image Processing

We wished to measure aging of the brain using three different brain morphological metrics, assessing different types of information and at different scales: (1) Voxel-wise GMD, a high-dimensional representation of the whole brain's gray matter, "close" to the acquired image; (2) Vertex-wise cortical thickness, a high-dimensional representation of the cortical gray matter; (3) Region-based cortical and subcortical volumes of interest, a low-dimensional representation of the gray matter of the brain.

Image processing was done using two pipelines, one for a voxel-based approach ("voxel-based morphometry (VBM)"; Ashburner and Friston 2000) and one for a surface-based approach (FreeSurfer; Fischl et al. 2002, 2004). An in-house

processing pipeline was used for the voxel-based approach (see, e.g., Nieuwenhuis et al. 2012): First, the T_1-weighted images were reoriented into Talairach orientation without scaling and corrected for magnetic field inhomogeneities (Sled et al. 1998). Intracranial masks were created and manually checked and edited if necessary, as described earlier for the baseline measurement. These were used as a basis to create intracranial masks at follow-ups (Peper et al. 2008; van Soelen et al. 2013). The intracranial volume (ICV) was then separated into gray matter, white matter, and cerebrospinal fluid using a partial volume segmentation algorithm (Brouwer et al. 2010). The gray matter fractions were blurred using an 8-mm full-width half-maximum kernel and resampled to a $2 \times 2 \times 2.4$ mm^3 resolution, resulting in so-called GMD maps. The T_1-weighted images were then warped by a combination of linear and nonlinear transformations to a model brain (Collins et al. 1995). The GMD images were warped into model space using these same transformations. For the second approach, Freesurfer version 5.3 was run with the T_1-weighted images in Talairach orientation as input. The manually edited intracranial masks were used to adjust the brain mask images after the initial processing steps. Global and subcortical volumes were extracted (Desikan–Killiany atlas; Desikan et al. 2006). We also extracted data on vertex level. We first averaged all scans in the study using the FreeSurfer subroutine fsaverage. Subsequently, to limit the number of vertices, we decimated the average template with a factor 10 using the subroutine mri_decimate and mapped all subjects to that template before extracting the data for each individual. Despite the availability of longitudinal data, we only used data from the cross-sectional pipeline for two reasons: One, subjects in our study had up to three scans included, and we wanted to avoid bias in segmentation accuracy between images. Two, for practical reasons, brain age should be computable from one image only. This approach has the consequence that we did not use the most accurate segmentation possible for participants that had longitudinal data available. For subsequent analyses, we expect the effect of this to be small, if anything, slightly increased noise levels result in an underestimation of heritability.

Features

We selected three brain feature sets to serve as input for the machine learning algorithm: (1) a voxel-based GMD feature set (157 256 features), (2) resampled (30 741 features) vertex-based cortical thickness values extracted from the FreeSurfer pipeline, (3) a feature set consisting of 90 local volumes extracted from the FreeSurfer pipeline (cortical volumes, subcortical, and cerebellar gray matter volumes, lateral, third, and fourth ventricle volumes). The measures of this third set were corrected for ICV by division, since global volume will be confounded by sex effects. These feature sets were chosen because we wanted to investigate brain-aging substrates with as much as possible variation in scale. We chose gray matter measures only because much of the white matter information is on the gray–white matter interface, which is also captured by gray matter. Before entering the training data into the algorithm, we removed features without variance and features for which more than 10% of the subjects had a value of 0. This only occurred in the vertex-wise analyses (6.1% of the features). Remaining zeros were imputed with the median value (0.07% of the data). Subsequently, the feature data were centered at zero, scaled to unit variance and scaling parameters were saved.

Brain Age Modeling

Three brain age models were built by training a support vector regression (SVR) machine (Drucker et al. 1997) to predict age as a weighted sum of the scaled features (plus an offset) from each of the three sets of brain imaging features (see previous section), according to the design described below. Our choice for SVR was based on its favorable properties with respect to robustness against high-dimensional feature sets, the possibility to incorporate linear and nonlinear kernels, and the good interpretability of the resulting linear models. We initially tested SVR models using both a linear and a radial basis function (RBF) kernel. Performance did not improve using the RBF kernel. Since interpretation of the weights of the linear model is easier, in the rest of this manuscript we describe results of the linear kernel only.

The data set was split into a training set and an application set. If a subject's age is estimated from his/her brain scan by applying a brain age model to that scan, none of this subject's brain scans nor any scan from his/her family members should be used in the training procedure of this model—otherwise the brain age estimate will be biased. The following modeling design, based on the leave-family-out (LOFO) procedure (Dubois et al. 2018), ensures unbiased brain age estimates.

The scans of the subjects were organized per family. For each of the 107 families, one scan from only one family member was selected for the training set. The training set thus consisted of 107 scans from 107 unrelated subjects (49 males, 58 females) from the full study, chosen such that the range and variance of the ages in the training set was as large as possible. Within the training set, we used nested cross-validation to build the brain age models: in each outer fold ($k = 10$), approximately 11% of the subjects were left out. The remaining training data were then used to select the optimal value for the cost-parameter C, using 5-fold cross-validation (inner folds) (FreeSurfer-based models only; voxel-based models used a previously optimized C; Schnack et al. 2016). The optimal C was defined by the value minimizing the root mean square error and was subsequently used to train the model on the training data excluding the left-out subjects. This final model was then used to predict the age of all the available scans of the subjects that were left out (i.e., the subjects in the outer folds and those of their siblings). We repeated this procedure 25 times to account for variation in the selection of the folds and averaged the 25 predicted ages for each scan. We used SVR as implemented in the R library film (https://bitbucket.org/RonaldJJ/film/src/master/) to build the brain age models.

To assess the performance of these models, we computed the mean absolute error (MAE)—the average deviation of the predicted age from the chronological age—and R^2, the proportion of chronological age variance explained by the model for all the scans in the training set. In what follows, we will call the predicted age the "brain age" and deviations from the chronological age will be the "brain age gap." After computing MAE and R^2, we regressed brain age against chronological age in the training set using the LOFO procedure, to remove the "regression to the mean effect" present in brain age estimates (Le et al. 2018). The coefficients of these regressions were used to compute corrected brain age gaps in the left-out subjects and their family members. This correction is a linear transformation of the data, similar for each subject, which aims to improve the estimate of the true age gap but does not change the relationship between subjects. It therefore should not influence the heritability and correlation

analyses that depend on these relationships, rather than on the absolute values. Indeed, when we reran analyses without age correction, the findings did not change and we only present results based on the corrected brain age gaps here.

Sex Differences

To assess the age-dependent sex differences in corrected brain age gaps, we used linear mixed modeling including fixed effects for age, sex, their interactions, and random effects for subjects nested within families. We fitted linear, quadratic, and cubic age effects in the sexes separately and proceeded with the model that had the lowest Akaike Information Criterion. In this model, we tested for significance of the effect of sex. Linear mixed modeling was performed using the lme4 package in R (Bates et al. 2015), using maximum likelihood estimation for model fitting as this allows for statistical comparisons of nested models.

We additionally created a male and female brain age model. We followed the LOFO procedure as described above and separated the remaining subjects in the training set on sex to train sex-specific models. We then predicted brain age in the left-out subjects and their siblings, regardless of their sex and repeated this procedure 25 times. For each subject, this resulted in a same-sex prediction, where as an example, a male's brain age was predicted on a males-trained model, and in an opposite sex prediction, where a male's brain age was predicted using a females-trained model. We computed MAE and R^2 for the training set for the male and female brain age prediction models and corrected brain age gaps in the full cohort, as described above.

Genetic Modeling

Twin and twin-sibling studies allow for disentanglement of additive genetic factors (*A*), common environmental factors (*C*), and unique environmental factors (*E*) (Boomsma et al. 2002; Posthuma and Boomsma 2000), by exploiting the fact that monozygotic (MZ) twins share almost 100% of their genetic makeup, whereas dizygotic (DZ) twins and sibling pairs share on average 50% of their aggregating genes. Heritability of the brain age gaps suggests that genetic factors influence an individual's deviation from expected average age trajectories. Given the mean effects of sex on the age trajectories, group mean age and sex effects were regressed out of the brain age gaps, thereby removing the developmental difference between a twin pair and their sibling, based on the best-fitting model identified above. Residuals of this regression were entered in the genetic model. First, we investigate the correlations between brain age gaps over time in a bivariate Cholesky model including *A*, *C*, and *E* (the ACE-model). We chose a generic model for all three time points for easier interpretation of results, even though some of the MZ/DZ correlations indicated an absence of common environmental influences (Supplementary Table S2). These phenotypic (observed) correlations (R_{ph}) can be separated into a genetic ($R_{ph\text{-}a}$) and environmental ($R_{ph\text{-}e}$) component based on cross-twin/cross-trait correlations. Significance of genetic overlap was established by determining whether the fit significantly deteriorated when the genetic correlation was constrained to zero. Likewise, when the genetic correlation could not be constrained to one (full overlap) without deteriorating the fit, we concluded that different genetic factors play a role at the different time points. This also indirectly tests for the necessity to include more than one genetic factor in the model. This will answer the question whether the genetic factors influencing brain age gaps are the same throughout development. Longitudinal twin models additionally allow for investigating the heritability of change in brain age gap (Teeuw et al. 2019; van Soelen et al. 2012a). A significant contribution of genetic factors on the change in brain age would suggest that the speed of the individual's developmental brain morphology pattern compared with his/her peers is driven by genetic factors. We implemented a trivariate Cholesky model including *A*, *C*, and *E* to estimate heritability at all three measurements to estimate heritability of brain age gap and brain age gap change (Supplementary Fig. S1). Finally, we investigate correlations between brain age gaps between the different modalities and tested whether the genetic correlation could be constrained to zero or one to assess whether the brain age gaps computed from the different feature sets are driven by the same or different genetic background in a series of bivariate ACE models.

Post Hoc Analyses

Brain age based on volumetric (ROI) measures was corrected for overall volume, which is necessary given the well-documented sex difference in brain size between males and females. For the same reason, we chose VBM GMD measures and not optimized VBM, as that method would also contain brain size information in the features. For cortical thickness, we initially chose uncorrected values, as cortical thickness is less related to overall brain size and sex differences in mean cortical thickness are limited in our cohort (Teeuw et al. 2019). It thus remains an open question whether brain age findings based on the cortical thickness feature set are driven by mean cortical thickness or regional variations in cortical thickness during development. To investigate this question, we repeated the analyses using brain age estimates based on corrected cortical thickness values (by division by mean cortical thickness) and brain age estimates based on mean cortical thickness alone (i.e., linear regression, using the LOFO procedure with 25 repeats as before).

Results

The model trained on the voxel-based GMD feature set outperformed those on volumetric ROI-based and vertex-based thickness feature sets (Fig. 1A), having both a smaller MAE (1.24 year) and higher R^2 (0.77). Over time, the brain age gaps were relatively stable, with significant correlations that ranged between 0.6 and 0.8 (Fig. 2). These correlations over time were similar for males and females (Supplementary Tables S3 and S4). Supplementary Table S5 contains the MAE and explained variance R^2 for brain age models based on the different feature sets.

(Sex Differences in) Brain Age (Gap): Evolution Over Time

Age gap development was described best with a cubic curve, with different age trajectories for males and females ($P < 0.001$) for all three of the feature sets (Fig. 1B). Males on average had a negative age gap (estimated to be younger than their chronological age) and females on average had a positive age gap (estimated to be older than their chronological age) in the combined model in a period from around 12 to 19 years of age.

We then trained brain age prediction models in the female and male subjects of the training set separately, using the same

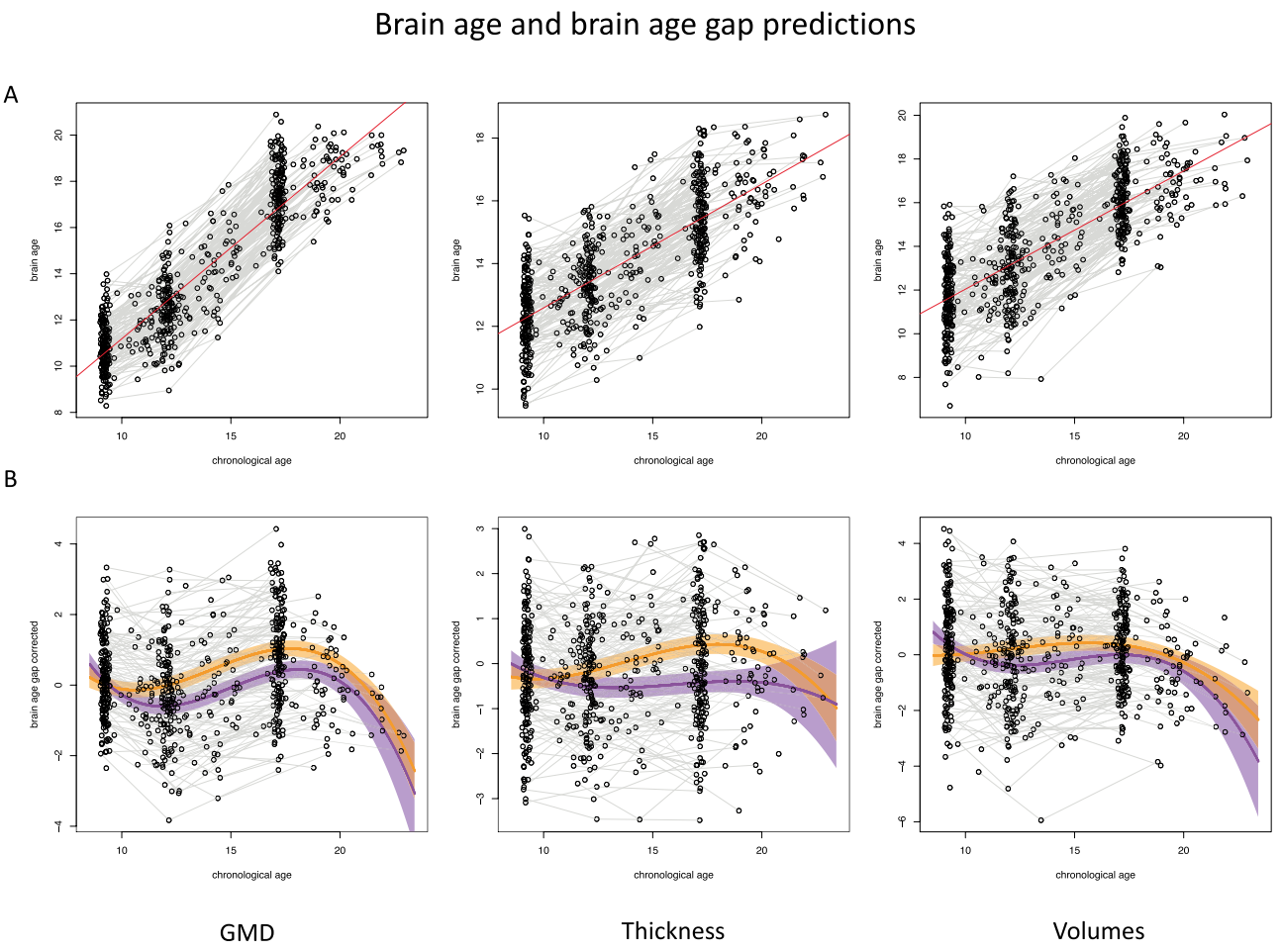

Figure 1. (A) Brain age predictions versus chronological age for models trained on GMD in voxels (left), vertex-wise thickness (middle), and volumes (ROIs, right) in the combined training set. (B) Regression-attenuation corrected brain age gap trajectories over time for females (in orange) and males (in purple). The plotted 95% confidence intervals are based on fixed effects uncertainty.

LOFO procedure as described above. For voxel-based models, the combined model outperformed both the female and the male model. For models trained on volumetric ROIs and vertex-wise cortical thickness, the female model performed as good as or better than the combined model, and the male model performed worse than the combined model (Supplementary Table S5). Using the opposite-sex predictions, that is, predicting the females' brain age gap based on the male model and vice versa, showed the same pattern of sex differences across our age range as in the combined model (Supplementary Fig. S2).

Figure 3 shows the weight maps for the volumetric ROI features, for the combined, female, and male models. The female model showed generally larger absolute weights than males (Supplementary Table S6). These ROI-based models are provided in the Appendix for application to other data sets.

Genetics of the Brain Age Gap and Its Evolution Over Time

Brain age gaps based on vertex-based thickness and voxel-based GMD were significantly heritable, with estimated heritability ranging between 56% and 79%. The heritability of volume (ROI)-based brain age gaps was significant around age 10 only (32%). The amount of change in GMD age gaps was significantly heritable (21% and 31% over two time periods, respectively). For thickness-based age gaps, the amount of change was significantly heritable in the second time interval only (19%). See Table 1 for heritability estimates and Supplementary Table S2 for twin and twin-sibling correlations. Making use of the cross-twin cross-time correlations, we found stable genetic factors to explain part of the correlations between the age gaps over time within the same feature set, for all three the feature sets (Fig. 2 and Supplementary Table S3).

Brain Age Gaps at Different Spatial Scales

Correlations between brain age gaps based on different feature sets were significant but substantially lower (0.15–0.40) than correlations over time within feature sets. Correlations between thickness-based and GMD-based brain age gap measures were also in part explained by shared genetic factors, but not fully: the genetic correlation between thickness- and GMD-based brain age gaps, specifically at age ~18, was significantly lower than 1, suggesting the existence of distinct genetic factors influencing the brain age measures separately (Fig. 2 and Supplementary Table S3).

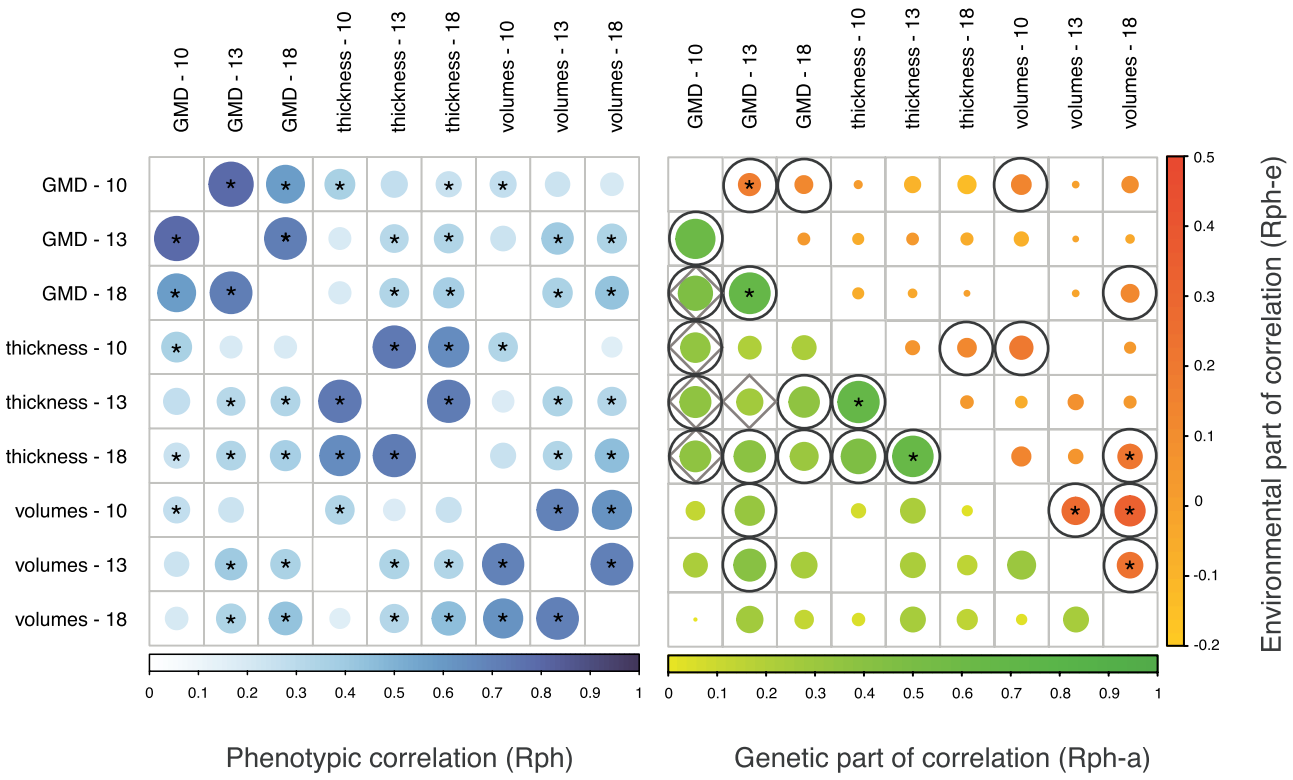

Figure 2. Significant phenotypic correlations over time and between different brain age gap estimates are depicted by blue circles. Circles are both scaled and colored according to the size of the correlations. Right: Correlations separated into a genetic ($R_{ph\text{-}a}$; lower diagonal in green colors) and unique environmental part ($R_{ph\text{-}e}$; upper diagonal in yellow–orange colors). The common environmental part ($R_{ph\text{-}c}$) was estimated to be small, nonsignificant, and was not displayed here. Black circles were added to correlations that were driven by a significant component that was shared by both phenotypes. Gray diamonds represent correlations for which the overlap was not complete, that is, the genetic/environmental correlation was significantly different from 1. In both panels, correlations that survived Bonferroni comparisons for multiple comparisons (0.05/72) are marked with ∗.

Table 1 Heritability of brain age gap*

Feature set	Variance components ($h^2/c^2/e^2$)				
	(age ~ 10)	(age ~ 13)	(age ~ 18)	(age ~ 10—> ~ 13)	(age ~ 13—> ~ 18)
GMD (Voxels)	h^2: **0.56** (0.32–0.71)	h^2: **0.64** (0.33–0.83)	h^2: **0.74** (0.53–0.85)	h^2: **0.21** (0.02–0.53)	h^2: **0.32** (0.03–0.64)
	c^2: 0.00 (0.00–0.14)	c^2: 0.06 (0.00–0.28)	c^2: 0.02 (0.00–0.18)	c^2: 0.09 (0.00–0.30)	c^2: 0.03 (0.00–0.19)
	e^2: 0.43 (0.28–0.62)	e^2: 0.30 (0.17–0.48)	e^2: 0.24 (0.15–0.38)	e^2: 0.70 (0.46–0.87)	e^2: 0.66 (0.35–0.90)
Cortical thickness (vertices)	h^2: **0.57** (0.35–0.75)	h^2: **0.79** (0.57–0.90)	h^2: **0.59** (0.28–0.76)	h^2: 0.10 (0.00–0.40)	h^2: **0.19** (0.02–0.51)
	c^2: 0.08 (0.00–0.26)	c^2: 0.00 (0.00–0.16)	c^2: 0.04 (0.00–0.28)	c^2: 0.14 (0.00–0.33)	c^2: 0.07 (0.00–0.25)
	e^2: 0.35 (0.26–0.51)	e^2: 0.21 (0.13–041)	e^2: 0.37 (0.26–0.56)	e^2: 0.76 (0.57–0.96)	e^2: 0.74 (0.40–0.99)
Volumes (ROI)	h^2: **0.32** (0.01–0.57)	h^2: 0.32 (0.00–0.61)	h^2: 0.15 (0.00–0.53)	h^2: 0.17 (0.00–0.39)	h^2: 0.17 (0.00–0.38)
	c^2: 0.13 (0.00–0.33)	c^2: 0.15 (0.00–0.38)	c^2: 0.29 (0.00–0.46)	c^2: 0.00 (0.00–0.21)	c^2: 0.00 (0.00–0.16)
	e^2: 0.55 (0.39–0.73)	e^2: 0.53 (0.34–0.76)	e^2: 0.56 (0.39–0.73)	e^2: 0.83 (0.61–1.00)	e^2: 0.83 (0.62–1.00)

*Brain age gaps were corrected for residual age effects based on the best-fitting model (cubic age effect, separately for the sexes). The variance was separated into an additive genetic component (h^2), a common environmental component (c^2), and a unique environmental component (e^2). Significant h^2 or c^2 components ($P < 0.05$) are displayed in bold.

Post Hoc Analyses—the Effect of Mean Cortical Thickness

We repeated the cortical thickness based brain age training our model on corrected cortical thickness values (by division by mean cortical thickness) and brain age estimates based on mean cortical thickness alone. The model based on corrected cortical thickness values performed slightly worse (MAE of 2.10 and R^2 of 0.41) and the model based on mean cortical thickness alone performed considerably worse (MAE of 2.63 and R^2 of 0.09) compared with the original model (c.f. MAE of 1.94 and R^2 of 0.49). Sex differences in the developmental trajectory were present and similar in the vertex-based corrected and uncorrected models ($P < 0.001$) but not significant for the brain age estimates based on mean cortical thickness alone ($P = 0.22$). In contrast, heritabilities of brain age gaps were similar for mean cortical thickness-based models (54%/63%/71%) but lower at

the second and third wave for vertex-based corrected cortical thickness values (60%/38%/48%; c.f. 57%/79%/59% in the original case).

Discussion

We investigated the brain structural development between 9 and 23 years of age using brain age, a marker that approximates the actual, neuroanatomical, age of the brain. The genetically informative longitudinal design of the study, with up to three scans per subject and its size (673 scans in total) allowed us to investigate possible sex-dependent speed of development, and the extent to which genes contribute to the developmental trajectory. Our main findings include on average a more advanced brain age observed in females compared with males during adolescence, for all three investigated brain age modalities: GMD, cortical thickness, and regional volumes. Individual brain age gaps were relatively stable over time, as inferred by high correlations of brain age gaps at different ages, implying that children that had higher brain age gaps early in adolescence were also still ahead 3 and 8 years later. Individual brain age was to a large extent determined by genetic factors with heritability estimates up to 79% for brain age gaps derived from local gray matter measures. The genetic overlap between cortical thickness-and GMD-derived brain age gap measures indicates that both represent the same kind of underlying biological brain age. Finally, for GMD brain age gaps, the change in brain age gap over time was also heritable, indicating that the speed of development for an individual was driven by genes.

Brain Age in Females and Males

At the beginning of puberty around age 9, we found very limited sex differences in brain age gap from the combined model. From around age 12, we observe a positive brain age gap in the females—representing an advanced estimated brain age—and a negative brain age gap in males, indicating a brain that is estimated to be younger than their chronological age. The maximum difference between males and females is about 1 year and occurs at ages 14–16, depending on the brain age model used. At age 18, brain age gaps in males and females seem to converge. This waxing and waning of an advanced brain age in females is in line with females reaching puberty earlier, evidenced by, for example, an earlier progression of Tanner status (Mul et al. 2001) and earlier increase of puberty-related hormone levels (Koenis et al. 2013). A maximal difference of 1 year might be a conservative estimate, since brains may develop differently in males and females, thereby confounding the brain age estimators. Interestingly, when we trained separate brain age models for males and females and applied these models to the opposite sex, age estimates were about as good and similar patterns of brain age gaps trajectories throughout adolescence were found as in the models that were trained on both sexes together. The same patterns of brain age gap changes with age were found. This suggests that the patterns of brain development do not differ very much between males and females and that differences in predicted age, or brain age, are due to differences in speed (or onset) of the same developmental processes. It should be noted, however, that this does not imply that there are no differences between males and females in brain structure. Stable differences, for instance, such as the well-documented approximately 10% larger brain in males compared with females (Ruigrok et al. 2014), will not be "seen" by age predictors. It must be noted that in both our ROI-based as well as our GMD-based models, this sex effect was taken out beforehand. The cortical thickness-based model did contain information on mean cortical thickness, but this did not drive the sex effect. Hence, differences between males and females in the developmental pattern of brain age gap were driven by local variations in cortical thickness.

The ROI-based models (Fig. 3) show the brain regions most contributing to the brain age estimates. The postcentral gyrus is among the top-ranked features for both the female and male model, which has been observed before as one of the contributors to brain age predictions (Khundrakpam et al. 2015). That study also showed a prominent role for the frontal association areas, which we do not find here. Instead, we observe that many of the top-ranked regions include the cingulate, medial orbitofrontal area, and nucleus accumbens, all part of the emotional limbic system, which plays a role in emotional processes and reward valuation (Rolls 2015). Given that adolescents show high emotional sensitivity and reward-seeking behavior (Casey et al. 2008), we may hypothesize that our brain age estimate represents emotional maturity.

Shared (Genetic) Influences on Brain Age and Brain Age Change

All brain age gaps correlated, with highest correlations found over time within modality.

We found a significant genetic overlap between brain age based on local GMD and based on local cortical thickness, thus indicating a shared genetic influence on the development of these brain properties. Brain age based on ROI volumes did not have significant genetic overlap with the other measures. This may indicate a developmental process with different genetic background, but it may also be due to the effect of averaging out over larger regions in the brain as higher spatial accuracy has been shown to provide better brain age estimates (Khundrakpam et al. 2015). Indeed, the precision (in terms of MAE) of the ROI-based model was substantially lower than the precision of the other two models. Finding genetic variants influencing brain age, such as was recently done in an adult cohort (Kaufmann et al. 2019) may help identify biological pathways that influence brain age, and potentially explain the brain age deficits that have been observed in psychiatric disorders (Hajek et al. 2019; Koutsouleris et al. 2014; Nenadić et al. 2017; Schnack et al. 2016). Apart from genetic influences on the brain age gaps themselves, we also found a significant genetic contribution to change in brain age gap over time. This indicates that individual differences in the speed of development are driven by genetic factors. This is important, since it has been recognized that longitudinal changes, rather than cross-sectional measurements, are more informative on the development of psychiatric disorders (Rapoport and Gogtay 2008; Shaw et al. 2010). We recently found genetic variants that influence longitudinal structural brain changes in an age-dependent manner (Brouwer et al. 2020). Focusing such studies on adolescent brain age specifically may aid in determining biological pathways associated with optimal development.

Different Modalities and Reliability of Brain Age

Supplementary Figure S3 shows a summary of the findings for the different brain age modalities. The combination of MAE of

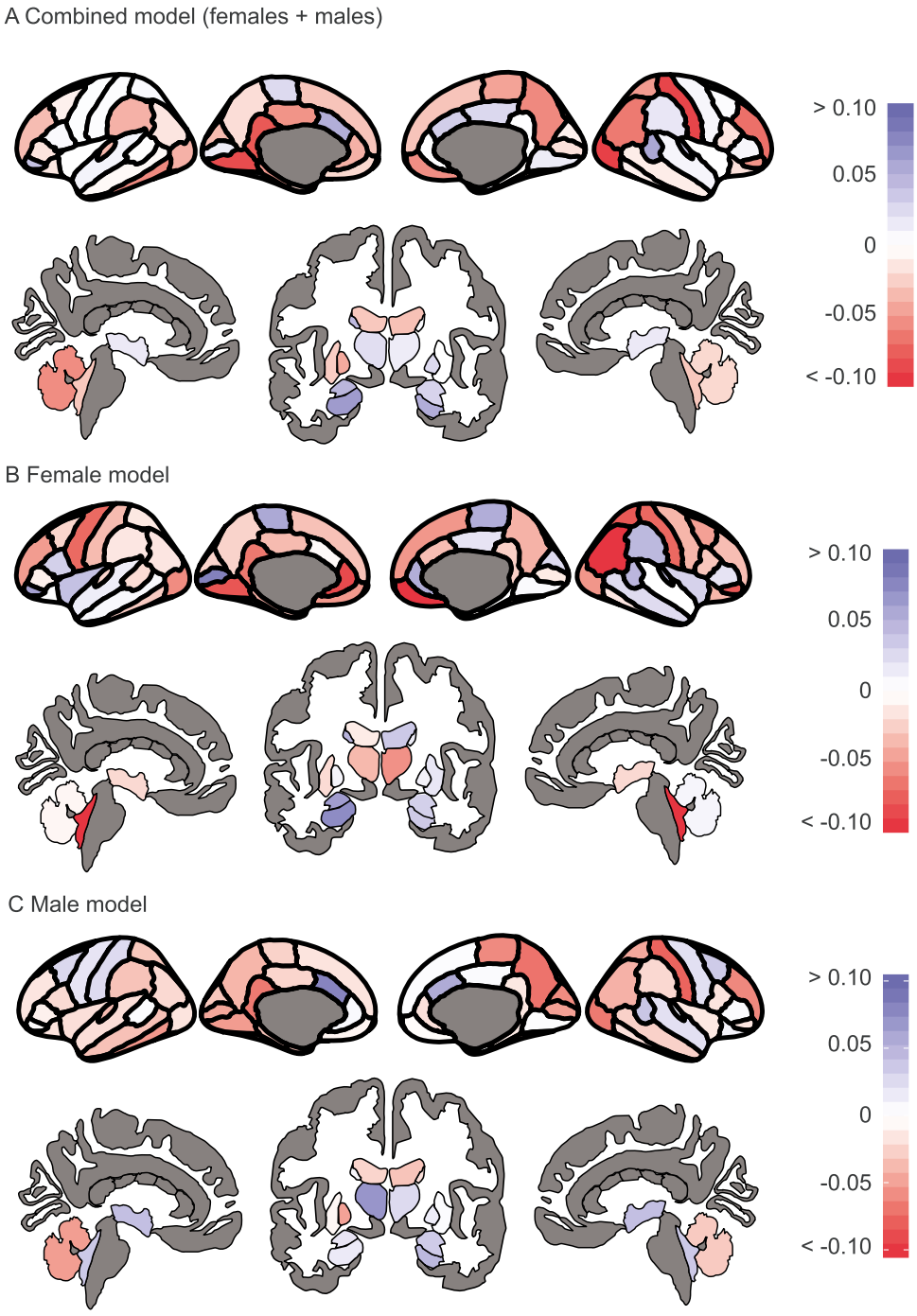

Figure 3. Feature weights of the ROI-based models: (*A*) the combined model, (*B*) the model based on the female training set, and (*C*) the male training set. Not shown are the feature weights for the L/R nucleus accumbens and L/R inferior lateral ventricle (0.024/−0.114 and 0.011/−0.011 for the combined model, −0.039/−0.102 and −0.068/−0.10 for female model, 0.017/−0.063, and 0.015/−0.011 for the male model, respectively).

prediction and the correlation of brain age gap between different time points can inform us about how accurate and how reliable the biological age of the brain is predicted. Voxel-wise GMD models showed low MAE and high over-time correlations, indicative of an accurate and reliable brain age prediction. ROI volume-based models had medium MAE but high over-time correlations, suggesting reliable, but somewhat biased predictions: the predicted age is likely a combination of true biological age and biological factors that are not related to age. Finally, the vertex-wise cortical thickness-based models showed high MAE and high over-time correlations, indicating reliable but even less accurate age predictions: the predicted biological age is biased due to relatively large contributions of biological factors not related to age. An example of biological factors not related to age could be brain regions that show an association with age, and for which size is also related to a function that has

influenced the size. If the size of such a volume is large because of its association with brain function, it could be interpreted as higher brain age by our models. Another example is that some individuals will have brain morphology that resembles the morphology of an older brain from the perspective of our age model, for example due to their genetic background, but which may not be informative for brain age for that subject. Longitudinal measures as well as different feature sets may help to disentangle these biological factors.

For the surface-based models, MAEs for the male models were higher than for the female models. This notion that female ages could be predicted more accurately than male ages suggests that there is less variation (in timing) of the developing female brain than there is in the male brain. Other developmental studies indeed have shown larger variation in male brain structure than in female brain structure (Forde et al. 2019; Wierenga et al. 2019; Wierenga et al. 2018). A more homogeneous sample gives rise to stronger relationships between brain features and age, which is reflected by larger regression coefficients, or feature weights. Indeed, the female models displayed larger feature weights than the male models. For the combined models, the situation is more complex due to the trade-off between heterogeneity and the positive effect of sample size on model performance. With respect to the female models, the increased heterogeneity of the combined sample resulted in less accurate models (higher MAE). With respect to the male models, incorporating relatively large heterogeneity, the effect of a twice as large sample outweighed the increase in heterogeneity: the combined models' MAE was lower.

Strengths and Limitations

Strengths of our study include the large number of scans, the longitudinal design (up to three scans per subject) and the fact that we measured twins and siblings, allowing for genetic analyses. There are also some limitations to be mentioned. Our age range started at 9 years; therefore, we cannot extrapolate our findings to years prior to this age. Furthermore, although the overall sample size is quite large, when split in males and females, training samples were modest in size. Future studies should thus aim at assessing brain age at earlier ages. Combining our models with those obtained in other (longitudinal) studies could validate and further refine the results.

Conclusions and Future Directions

To conclude, we have built a reliable brain age model for adolescence (ages 9–23) and applied that to the longitudinal MRI scans of twin subjects and their siblings. Brain development ("aging") in males and females follows comparable patterns, but in females group level brain age is at most 1 year ahead of group level brain age in males. At the individual level, being ahead or not is a heritable phenotype. In future studies, advanced/delayed brain aging may be related to neuropsychological or behavioral traits.

Supplementary Material

Supplementary material can be found at Cerebral *Cortex* online.

Funding

The Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO 51.02.061 to H.H., NWO 51.02.062 to D.B., NWO-NIHC Programs of excellence 433-09-220 to H.H., NWO-MagW 480-04-004 to D.B., and NWO/SPI 56-464-14192 to D.B.); FP7 Ideas: European Research Council (ERC-230374 to D.B.); and Universiteit Utrecht (High Potential Grant to H.H.).

Notes

We thank all participants and their families. *Conflict of Interest*: None declared.

Appendix

FreeSurfer ROI-based model for brain age estimation in adolescents. https://github.com/patterns-in-psychiatry/abam.

References

Ashburner J, Friston KJ. 2000. Voxel-based morphometry - the methods. *Neuroimage*. 11(6 Pt 1):805–821. https://doi.org/10.1006/nimg.2000.0582.

Bates D, Mächler M, Bolker BM, Walker SC. 2015. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 67(1). https://doi.org/10.18637/jss.v067.i01.

Boomsma D, Busjahn A, Peltonen L. 2002. Classical twin studies and beyond. *Nat Rev Genet*. 3(11):872–882. https://doi.org/10.1038/nrg932.

Boomsma DI, De Geus EJC, Vink JM, Stubbe JH, Distel MA, Hottenga JJ, Posthuma D, van Beijsterveldt TC, Hudziak JJ, Bartels M *et al*. 2006. Netherlands twin register: from twins to twin families. *Twin Res Hum Gen*. 9(6):849–857. https://doi.org/10.1375/183242706779462426.

Brouwer RM, Hulshoff Pol HE, Schnack HG. 2010. Segmentation of MRI brain scans using non-uniform partial volume densities. *Neuroimage*. 49(1):467–477. https://doi.org/10.1016/j.neuroimage.2009.07.041.

Brouwer RM, Klein M, Grasby KL, Schnack HG, Jahanshad N, Teeuw J, Thomopoulos SI, Sprooten E, Franz CE, Gogtay N *et al*., 2020. Dynamics of brain structure and its genetic architecture over the lifespan. *BioRxiv*. https://doi.org/10.1101/2020.04.24.031138.

Brown TT, Kuperman JM, Chung Y, Erhart M, McCabe C, Hagler DJ Jr, Venkatraman VK, Akshoomoff N, Amaral DG, Bloss CS *et al*. 2012. Neuroanatomical assessment of biological maturity. *Curr Biol*. 22(18):1693–1698. https://doi.org/10.1016/j.cub.2012.07.002.

Cao B, Mwangi B, Hasan KM, Selvaraj S, Zeni CP, Zunta-Soares GB, Soares JC. 2015. Development and validation of a brain maturation index using longitudinal neuroanatomical scans. *Neuroimage* 117:311–318. https://doi.org/10.1016/j.neuroimage.2015.05.071.

Casey BJ, Jones RM, Hare TA. 2008. The adolescent brain. *Ann N Y Acad Sci*. 124:111–126. https://doi.org/10.1196/annals.1440.010.

Cole JH, Poudel RPK, Tsagkrasoulis D, Caan MWA, Steves C, Spector TD, Montana G. 2017. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *Neuroimage*. 163:115–124. https://doi.org/10.1016/j.neuroimage.2017.07.059.

Collins DL, Holmes CJ, Peters TM, Evans AC. 1995. Automatic 3-D model-based neuroanatomical segmentation. *Hum Brain Mapp*. 3(3):190–208. https://doi.org/10.1002/hbm.460030304.

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT *et al*. 2006. An automated labeling system for subdividing the human

cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 31(3):968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021.

Dosenbach NUF, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN *et al*. 2010. Prediction of individual brain maturity using fMRI. *Science*. 329(5997):1358–1361. https://doi.org/10.1126/science.1194144.

Drucker H, Surges CJC, Kaufman L, Smola A, Vapnik V. 1997. Support vector regression machines. *Adv Neural Inform Process Syst*. 9:155–161.

Dubois J, Galdi P, Paul LK, Adolphs R. 2018. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philos Trans R Soc B Biol Sci*. 373(1756). https://doi.org/10.1098/rstb.2017.0284.

Erus G, Battapady H, Satterthwaite TD, Hakonarson H, Gur RE, Davatzikos C, Gur RC. 2015. Imaging patterns of brain development and their relationship to cognition. *Cereb Cortex*. 25(6):1676–1684. https://doi.org/10.1093/cercor/bht425.

Fischl B, Salat D, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S *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, Salat DH, van der Kouwe AJ, Makris N, Segonne F, Quinn BT, Dale AM. 2004. Sequence-independent segmentation of magnetic resonance images. *Neuroimage*. 23(Suppl 1):S69–S84. https://doi.org/10.1016/j.neuroimage.2004.07.016.

Forde NJ, Jeyachandra J, Joseph M, Jacobs GR, Dickie R, Satterthwaite TD, Shinohara RT, Ameis SH, Voineskos A. 2019. Sex differences in variability of brain structure across the lifespan authors. *BioRxiv*. http://dx.doi.org/10.1101/842567.

Franke K, Ziegler G, Klöppel S, Gaser C. 2010. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. *Neuroimage*. 50(3):883–892. https://doi.org/10.1016/j.neuroimage.2010.01.005.

Franke K, Luders E, May A, Wilke M, Gaser C. 2012. Brain maturation: predicting individual BrainAGE in children and adolescents using structural MRI. *Neuroimage*. 63(3):1305–1312. https://doi.org/10.1016/j.neuroimage.2012.08.001.

Gaser C, Franke K, Kloppel S, Koutsouleris N, Sauer H, Alzheimer's Disease Neuroimaging Initiative. 2013. BrainAGE in mild cognitive impaired patients: predicting the conversion to Alzheimer's disease. *PLoS One*. 8(6):e67346. https://doi.org/10.1371/journal.pone.0067346.

Gennatas ED, Avants BB, Wolf DH, Satterthwaite TD, Ruparel K, Ciric R, Hakonsarson H, Gur RE, Gur RC 2017. Age-related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood. *J Neurosci*. 37(20):5065–5073. https://doi.org/10.1523/JNEUROSCI.3550-16.2017.

Giedd JN, Blumenthal JD, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, Evans AC, Paus T, Rapoport JL. 1999. Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci*. 2(10):861–863. https://doi.org/10.1038/13158.

Hajek T, Franke K, Kolenic M, Capkova J, Matejka M, Propper L, Uher R, Stopkova P, Novak T, Paus T *et al*. 2019. Brain age in early stages of bipolar disorders or schizophrenia. *Schizophr Bull*. 45(1):190–198. https://doi.org/10.1093/schbul/sbx172.

Herting MM, Sowell ER. 2017. Puberty and structural brain development in humans. *Front Neuroendocrinol*. 44:122–137. https://doi:10.1016/j.yfrne.2016.12.003.

Koenis MM, Brouwer RM, van Baal GC, van Soelen IL, Peper JS, van Leeuwen M, Delemarre-van de Waal HA, Boomsma DI, Hulshoff Pol HE. 2013. Longitudinal study of hormonal and physical development in young twins. *J Clin Endocrinol Metab*. 98(3):E518–E527. https://doi.org/10.1210/jc.2012-3361.

Kaufmann T, van der Meer D, Doan NT, Schwartz E, Lund MJ, Agartz I, Alnæs D, Barch DM, Baur-Streubel R, Bertolino A *et al*. 2019. Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature*. 22(10):1617–1623. https://doi.org/10.1038/s41593-019-0471-7

Koenis MMG, Brouwer RM, Swagerman SC, van Soelen ILC, Boomsma DI, Hulshoff Pol HE. 2018. Association between structural brain network efficiency and intelligence increases during adolescence. *Hum Brain Mapp*. 39(2):822–836. https://doi.org/10.1002/hbm.23885.

Koolschijn PCMP, Crone EA. 2013. Sex differences and structural brain maturation from childhood to early adulthood. *Dev Cogn Neurosci*. 5:106–118. https://doi.org/10.1016/j.dcn.2013.02.003.

Koutsouleris N, Davatzikos C, Borgwardt S, Gaser C, Bottlender R, Frodl T, Falkai P, Riecher-Rössler A, Möller HJ, Reiser M *et al*. 2014. Accelerated brain aging in schizophrenia and beyond: a neuroanatomical marker of psychiatric disorders. *Schizophr Bull*. 40(5):1140–1153. https://doi.org/10.1093/schbul/sbt142.

Khundrakpam BS, Tohka J, Evans AC, Brain Development Cooperative Group. 2015. Prediction of brain maturity based on cortical thickness at different spatial resolutions. *Neuroimage*. 111:350–359 https://doi.org/10.1016/j.neuroimage.2015.02.046.

Le TT, Kuplicki RT, McKinney BA, Yeh HW, Thompson WK, Paulus MP. 2018. A nonlinear simulation framework supports adjusting for age when analyzing BrainAGE. *Front Aging Neurosci*. 10(October):1–11 https://doi.org/10.3389/fnagi.2018.00317.

Lowe LC, Gaser C, Franke K, Alzheimer's Disease Neuroimaging Initiative. 2016. The effect of the APOE genotype on individual BrainAGE in normal aging, mild cognitive impairment, and Alzheimer's disease. *PLoS One*. 11(7):e0157514. https://doi.org/10.1371/journal.pone.0157514.

Mul D, Fredriks AM, van Buuren S, Oostdijk W, Verloove-Vanhorick SP, Wit JM. 2001. Pubertal development in the Netherlands 1965–1997. *Pediatr Res*. 50(4):479–486. https://doi.org/10.1203/00006450-200110000-00010.

Nenadić I, Dietzek M, Langbein K, Sauer H, Gaser C. 2017. BrainAGE score indicates accelerated brain aging in schizophrenia, but not bipolar disorder. *Psychiatry Res Neuroimaging*. . doi: https://doi.org/10.1016/j.pscychresns.2017.05.006.

Nieuwenhuis M, van Haren NEM, Hulshoff Pol HE, Cahn W, Kahn RS, Schnack HG. 2012. Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *Neuroimage*. . https://doi.org/10.1016/j.neuroimage.2012.03.079.

Peper JS, Brouwer RM, Schnack HG, van Baal GC., van Leeuwen M, van den Berg SM, Delemarre-van de Waal HA, Janke AL, Collins DL, Evans AC. 2008. Cerebral white matter in early puberty is associated with luteinizing hormone concentrations. *Psychoneuroendocrinology*. 33(7):909–915. https://doi.org/10.1016/j.psyneuen.2008.03.017

Peper JS, Brouwer RM, Schnack HG, van Baal GC, van Leeuwen M, van den Berg SM, Delemarre-van de Waal HA, Boomsma DI, Kahn RS, Hulshoff Pol HE. 2009. Sex steroids and brain structure in pubertal boys and girls.

Psychoneuroendocrinology. 34(3):332–342. https://doi.org/10.1016/j.psyneuen.2008.09.012.

Posthuma D, Boomsma DI. 2000. A note on the statistical Power in extended twin designs. *Behav Genet*. 30(2):147–158 https://doi.org/10.1023/A:1001959306025.

Rapoport JL, Gogtay N. 2008. Brain neuroplasticity in healthy, hyperactive and psychotic children: insights from neuroimaging. *Neuropsychopharmacology*. 33:181–197. https://doi.org/10.1038/sj.npp.1301553.

Rolls ET. 2015. Limbic systems for emotion and for memory, but no single limbic system. *Cortex*. 62:119–157. https://doi.org/10.1016/j.cortex.2013.12.005.

Ruigrok ANV, Salimi-Khorshidi G, Lai MC, Baron-Cohen S, Lombardo MV, Tait RJ, Suckling J. 2014. A meta-analysis of sex differences in human brain structure. *Neurosci Biobehav Rev*. 39:34–50. https://doi.org/10.1016/j.neubiorev.2013.12.004.

Schnack HG, van Haren NE, Nieuwenhuis M, Hulshoff Pol HE, Cahn W, Kahn RS. 2016. Accelerated brain aging in schizophrenia: a longitudinal pattern recognition study. *Am J Psychiatry*. 173(6):607–616. doi: 10.1176/appi.ajp.2015.15070922.

Shaw P, Gogtay N, Rapoport J. 2010. Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. *Hum Brain Mapp*. 31:917–925. doi: 10.1002/hbm.21028.

Sled JG, Zijdenbos AP, Evans AC. 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 17(1):87–97. doi: 10.1002/hbm.460030304.

Teeuw J, Brouwer RM, Koenis MMG, Swagerman SC, Boomsma DI, Hulshoff Pol HE. 2019. Genetic influences on the development of cerebral cortical thickness during childhood and adolescence in a Dutch longitudinal twin sample: the Brainscale Study. *Cereb Cortex*. 29(3):978–993. doi: 10.1093/cercor/bhy005.

Truelove-Hill M, Erus G, Bashyam V, Varol E, Sako C, Gur RC, Gur RE, Koutsouleris N, Zhuo C, Fan Y *et al*. 2020. 2020. A multidimensional neural maturation index reveals reproducible developmental patterns in children and adolescents. *J Neurosci*. 40(6):1265–1275. doi: 10.1523/JNEUROSCI.2092-19.2019.

Van Beijsterveldt CEM, Groen-Blokhuis M, Hottenga JJ, Franić S, Hudziak JJ, Lamb D, Huppertz C, de Zeeuw E, Nivard M, Schutte N *et al*. 2013. The young Netherlands twin register (YNTR): longitudinal twin and family studies in over 70,000 children. *Twin Res Hum Gen*. 16(1):252–267. doi: 10.1017/thg.2012.118.

van Soelen IL, Brouwer RM, van Baal GC, Schnack HG, Peper JS, Collins DL, Evans AC, Kahn RS, Boomsma DI, Hulshoff Pol HE. 2012a. Genetic influences on thinning of the cerebral cortex during development. *Neuroimage*. 59(4):3871–3880. doi: 10.1016/j.neuroimage.2011.11.044.

van Soelen IL, Brouwer RM, Peper JS, van Leeuwen M, Koenis MM, van Beijsterveldt TC, Swagerman SC, Kahn RS, Hulshoff Pol HE, Boomsma DI. 2012b. Brain SCALE: brain structure and cognition: an adolescent longitudinal twin study into the genetic etiology of individual differences. *Twin Res Hum Genet*. 15(3):453–467. doi: 10.1017/thg.2012.4.

van Soelen IL, Brouwer RM, van Baal GC, Schnack HG, Peper JS, Chen L, Kahn RS, Boomsma DI, Hulshoff Pol HE. 2013. Heritability of volumetric brain changes and height in children entering puberty. *Hum Brain Mapp*. 34(3):713–725. doi: 10.1002/hbm.21468.

Vandenbosch MMLJZ, van't Ent D, Boomsma DI, Anokhin AP, Smit DJA. 2019. EEG-based age-prediction models as stable and heritable indicators of brain maturational level in children and adolescents. *Hum Brain Mapp*. 40(6):1919–1926. doi: 10.1002/hbm.24501.

Vijayakumar N, Op de Macks Z, Shirtcliff EA, Pfeifer JH. 2018. Puberty and the human brain: insights into adolescent development. *Neurosci Biobehav Rev*. 92:417–436. doi: 10.1016/j.neubiorev.2018.06.004.

Wierenga LM, Sexton JA, Laake P, Giedd JN, Tamnes CK. 2018. A key characteristic of sex differences in the developing brain: greater variability in brain structure of boys than girls. *Cereb Cortex*. 28(8):2741–2751. doi: 10.1093/cercor/bhx154.

Wierenga LM, Bos MGN, van Rossenberg F, Crone EA. 2019. Sex effects on development of brain structure and executive functions: greater variance than mean effects. *J Cogn Neurosci*. 31(5):730–753. doi: 10.1162/jocn.