Prediction of Brain Maturity based on Cortical Thickness at Different Spatial Resolutions

Budhachandra S. Khundrakpam, Jussi Tohka, Alan C. Evans

PII: S1053-8119(15)00147-0
DOI: doi: 10.1016/j.neuroimage.2015.02.046
Reference: YNIMG 12017

To appear in: *NeuroImage*

Accepted date: 19 February 2015

Please cite this article as: Khundrakpam, Budhachandra S., Tohka, Jussi, Evans, Alan C., Prediction of Brain Maturity based on Cortical Thickness at Different Spatial Resolutions, *NeuroImage* (2015), doi: 10.1016/j.neuroimage.2015.02.046

Prediction of Brain Maturity based on Cortical Thickness at Different Spatial Resolutions

Budhachandra S. Khundrakpam[1‡*], Jussi Tohka[2‡], Alan C. Evans[1] and Brain Development Cooperative Group[†]

[1]Montreal Neurological Institute, McGill University, Montreal, Canada

[2]Department of Bioengineering and Aerospace Engineering, Universidad Carlos III de Madrid, Spain

[†]See Appendix for author list and affiliations of the Brain Development Cooperative Group

[‡]Equal authors

*** Corresponding Author:**

Budhachandra Khundrakpam

Montreal Neurological Institute, McGill University,

Montreal, Quebec, Canada - H3A 2B4

Email: budha@bic.mni.mcgill.ca

Phone: 5143986174

Number of words: 3810 [Abstract=247; Intro=634; Materials=1656; Results=752; Discussion=2041]

Number of Figures: 3

Number of Tables: 5

Abstract:

Several studies using magnetic resonance imaging (MRI) scans have shown developmental trajectories of cortical thickness. Cognitive milestones happen concurrently with these structural changes, and a delay in such changes has been implicated in developmental disorders such as attention-deficit/hyperactivity disorder (ADHD). Accurate estimation of individuals' brain maturity, therefore, is critical in establishing a baseline for normal brain development against which neurodevelopmental disorders can be assessed. In this study, cortical thickness derived from structural magnetic resonance imaging (MRI) scans of a large longitudinal dataset of normally growing children and adolescents (n = 308), were used to build a highly accurate predictive model for estimating chronological age (cross-validated correlation up to R = 0.84). Unlike previous studies which used kernelized approach in building prediction models, we used an elastic net penalized linear regression model capable of producing a spatially sparse, yet accurate predictive model of chronological age. Upon investigating different scales of cortical parcellation from 78 to 10240 brain parcels, we observed that the accuracy in estimated age improved with increased spatial scale of brain parcellation, with the best estimations obtained for spatial resolutions consisting of 2560 and 10240 brain parcels. The top predictors of brain maturity were found in highly localized sensorimotor and association areas. The results of our study demonstrate that cortical thickness can be used to estimate individuals' brain maturity with high accuracy, and the estimated ages relate to functional and behavioural measures, underscoring the relevance and scope of the study in the understanding of biological maturity.

Key Words:

Cortical thickness, prediction model, brain maturation, structural magnetic resonance imaging, elastic-net regularized regression

Introduction:

Comprehensive investigation of the maturational trajectories of brain structure have been facilitated by the advent of advanced magnetic resonance imaging (MRI) methods. MRI scans of normally developing children and adolescents have demonstrated developmental trajectories of gray matter (GM) volumes and cortical thickness (Giedd et al., 1999; Giedd and Rapoport, 2010; Gogtay et al., 2004; Shaw et al., 2008). Deviations in these normal brain developmental trajectories have been proposed to give rise to neurodevelopmental disorders such as ADHD (Paus et al. 2008; Shaw et al., 2007, 2010). As such, a single integrated reference curve for brain maturation might be useful for early diagnosis of neuropsychiatric disorders. Towards realizing this goal, recent studies have used multivariate machine learning algorithms to derive brain maturity curves (Brown et al., 2012; Dosenbach et al., 2010; Erus et al., 2014; Franke et al., 2012; Mwangi et al., 2013). While Franke et al. (2012) used a voxel based morphometry (VBM) approach of T1-weighted MRI scans to predict biological age with high accuracy (subjects with age ranging from 5 to 18 years, $R = 0.93$, mean absolute error, MAE = 1.1 years); Dosenbach et al. (2010) used resting state functional connectivity MRI to predict chronological age (subjects with age ranging from 7 to 30 years, $R^2 = 0.55$); and Erus et al. (2013) used diffusion tensor imaging (DTI)-based metrics of fractional anisotropy and diffusivity to predict individual subject's chronological age (subjects with age ranging from 4 to 85 years, $R = 0.89$). Brown et al. (2012) combined multiple imaging indices: T1-, T2- and diffusion-weighted imaging (subjects with age ranging from 3 to 20 years, $R = 0.96$, MAE = 1.0 years) while Erus et al. (2014) used T1-based regional volumetric maps of GM, WM and lateral ventricle to predict brain maturity (subjects with age ranging from 8 to 22 years, $R = 0.89$).

None of the earlier studies except Brown et al. (2012), have explored the contribution of cortical thickness in estimating chronological age. The particular study, however, used cortical thickness averaged over the whole brain and over the entire hemisphere (Brown et al., 2012). Such an averaging approach may miss important information as to which local cortical regions are top predictors of maturation. This is particularly relevant since localized changes in cortical thickness have proven to be sensitive indices of brain maturation in typical and atypical brain development (Ameis et al. 2014; Gogtay et al., 2004; Raznahan et al., 2010, 2011; Sharda et al. 2014; Shaw et al., 2008, 2012, 2013).

Thus, the principal aim of the study was to assess the top predictors of brain maturity based on cortical thickness. Additionally, since brain maturation and cognitive development have been shown to be associated with changes in cortical thickness in highly localized brain regions (Shaw et al., 2006), we aim to investigate the top predictors of brain maturity at high spatial resolutions. To realize this goal, we apply an elastic net penalized linear regression model that is capable of producing a spatially sparse, but yet accurate predictive model of chronological age. Our machine learning approach is different from kernelized approaches (either relying on support vector machines or relevance vector machines) of the previous studies that promote sparsity in the kernel space (Dosenbach et al., 2010; Erus et al., 2014; Franke et al., 2012; Mwangi et al., 2013). Enforcing sparsity in the kernel space does not ensure sparsity in data space, and thus the predictive models of the earlier studies have been spatially dense, meaning that most brain voxels contribute to the prediction models resulting to findings that are hard to interpret. Instead, by imposing a sparsity requirement directly on the data space, we obtained predictive models that are spatially sparse (few voxels or surface points contribute to prediction) and perhaps easier to interpret. A more technical account on this difference can be found in (Li et al., 2005).

Materials and Methods:

Participants

The data for the study were obtained from the Pediatric MRI Data Repository created for the NIH MRI Study of Normal Brain Development (Evans, 2006); a multi-site project providing a normative database to characterize healthy brain maturation in relation to behavior. Demographic details of the subjects used in the study are given in Table 1.

Insert **Table 1**

MRI Acquisition Protocol

For each participant, a three-dimensional T1-weighted Spoiled Gradient Recalled (SPGR) echo sequence suing 1.5 Tesla scanners was obtained, with 1mm isotropic data acquired sagittally from the entire head. Due to the limit of 124 slices in GE scanners, slice thickness of ~1.5 mm was acquired. Additionally, using a two-dimensional (2D) multi-slice (2mm) dual echo fast spin echo (FSE) sequence, T2-weighted (T2W) and proton density-weighted (PDW) images were acquired. The total acquisition time was about 25 minutes, and was often repeated when indicated by the scanner-side quality control process. Subjects which were not able to tolerate this procedure, received a fallback protocol that involved shorter 2D acquisitions with slice thickness of 3mm (Evans, 2006).

Cortical Thickness Measurements

All MRI images were processed using the CIVET pipeline developed at the MNI for fully automated structural image analysis (http://www.bic.mni.mcgill.ca/ServicesSoftware/CIVET).

The native MRI images were first corrected for non-uniformity artifacts using the N3 algorithms (Sled et al., 1998), and registered into the stereotaxic space (Talairach and Tournoux, 1988) using a 9-parameter linear transformation (Collins et al., 1994). The registered and corrected images were further segmented into gray matter, white matter, cerebrospinal fluid and background using an advanced neural net classifier (Zijdenbos et al., 2002), and fractional tissue content in each voxel was estimated (Tohka et al., 2004). Then, using the CLASP algorithm (Kabani et al., 2001; Kim et al., 2005; Lee et al., 2006; MacDonald et al., 2000), the inner and outer gray matter surfaces were automatically extracted from each MR volume. Lastly, cortical thickness was measured in native space using the linked distance between the two inner and outer gray matter surfaces at 81,924 vertices (163840 polygons) throughout the cortex (Lerch and Evans, 2005). A stringent quality control (QC) procedure was followed at several data pre-processing steps in order to make sure that there were no motion, surface-surface intersections, blood vessels, etc. (see Supplementary Table S1), resulting to the longitudinal data (679 scans) from 308 subjects for the study.

Age Prediction

We assumed a linear model for predicting subject's age based on cortical thickness measurements. The model is

$$AGE = \sum_{i=1}^{p} b_i T_i + K + \varepsilon \quad (1)$$

where *AGE* is the age of the subject (in days); T_i, $i = 1, \ldots, p$, are the cortical thickness measurements at the point i of p vertices; b_i and K are the model parameters, and ε is an error term. Before proceeding, we standardize the variables T_i so that each of them has unit variance

and zero mean. We denote these standardized thickness measurements for subject *i* at the point *j* by x_{ij}.

We considered several spatial resolutions of measurements for cortical thickness. The original 81,924 measurements on the cortical surface were grouped into smaller sets and averaged. The number of parcels chosen (*p*) were 78, 160, 640, 2560 and 10240. The case *p* = 78 was obtained by averaging cortical measurements in each cortical region of the Automated Anatomical Labelling (AAL) atlas (Tzourio-Mazoyer et al., 2002). The cases *p* = 160, 640, 2560 and 10240 were obtained by recursively merging the neighbouring triangles of the triangular surface mesh model.

The prediction models contained up to 10241 model parameters for data from *N* = 679 scans (308 subjects). This rendered the ordinary least squares (OLS)-based parameter estimation ill-posed. Therefore, we used penalized least squares approach with elastic net penalty (Zou and Hastie, 2005). This approach leads to simultaneous model parameter estimation and variable selection by forcing many parameters to zero value. We denote the standardized measurements for the subject *i* by $x_i = [x_{i1}, \dots, x_{ip}]^T$, and the model parameters by $\beta = [b_1, \dots, b_p]^T$. We aim to minimize the elastic net cost function, which is written as (Zou and Hastie, 2005) –

$$\frac{1}{2N}\sum_{i=1}^{N}(AGE_i - K - x_i^T\beta)^2 \; + \; \lambda\sum_{j=1}^{p}\left(\alpha\left|b_j\right| + 0.5(1-\alpha)(b_j)^2\right) \qquad (2)$$

Symbols α and λ in Eq. (2) denote regularization parameters and the latter term is the elastic net penalty. The elastic net penalty is a weighted sum of (i) the LASSO penalty $||\beta||_1 = \sum_{j=1}^{p}\left(\left|b_j\right|\right)$ and (ii) ridge regression penalty $0.5||\beta||^2 = \sum_{j=1}^{p}\left(0.5(b_j)^2\right)$. Due to L1-norm regularization, LASSO penalty forces many parameters to have zero values leading to sparse prediction models. While the model sparsity is a desired property, LASSO penalty will typically pick just one of

many correlated variables to be included in the model. This is not what is usually desired in neuroimaging applications (see, e.g. Carroll et al., 2009). Ridge regression penalty shrinks the coefficients of correlated predictors towards each other encouraging the model parameters for the correlated variables to have approximately the same value (this is often termed as grouping effect). However, using only the ridge penalty by setting $\alpha = 0$ in Eq. (2) would not result in 'the desired' sparse model with many zero model parameters. The elastic net is able to produce sparse models combined with the grouping effect and therefore provides a suitable regularization approach for variable selection in neuroimaging applications. We fixed the value α at 0.5 to give equal weight for the LASSO and ridge penalties. The parameter λ was selected by cross-validation as detailed below. This cost was then minimized using GLMNET algorithm (Friedman et al., 2010) with respect to β to find the model parameters for age estimation.

We selected λ and evaluated the age predictions using two nested stratified 10-fold cross-validation (CV) loops (Ambroise and McLachlan, 2002; Huttunen et al., 2012). The value for the parameter λ was selected in the inner CV loop and the age predictions were evaluated in the outer CV loop thus avoiding the training on testing data problem. The model goodness criteria applied were the correlation coefficient between the chronological and estimated age, and the mean absolute error (MAE) between the chronological and estimated age. The stratification of the CV means here that the age distribution of the subjects in each of the 10 CV folds was approximately the same. Since we had up to three cortical thickness measurements of certain subjects at different ages, we controlled for the non-independence of the observations in the CV in such a way that all the scans of the subject *i* were either in the training set or in the test set. This was done by ensuring that all the measurements of the same subject were in the same CV fold. Therefore, data from subject *i* was never used to build an age-predictor for subject *i* and we

could use all the data of 679 morphometric measurements from 308 subjects. This is an important consideration and failure to account for non-independence would lead to positively biased error model goodness criteria values.

We subjected the MAEs (between the estimated and chronological ages of the subjects) with different measurement resolutions to statistical hypothesis testing. The test applied was a permutation test between two sets of absolute values of subject-wise prediction errors and we were interested in testing whether predictions using a finer measurement scale (of 10240 cortical parcels) were more accurate than by using lower measurement resolutions (2560, 640, 160, or 78 cortical parcels). Because we standardized the measurements, the magnitude of the i^{th} coefficient $|b_j|$ provides a measure of the relative strength of i^{th} predictor and thus, cortical parcels s can be ranked based on the regression coefficient magnitudes.

Next, we explored the effect of gender and acquisition site on the performance of the prediction model. For gender effect, we trained two different age-models, one for boys and one for girls, and compared the accuracy of the estimated ages derived this way to the accuracy of the estimated ages derived using the combined model (that includes both boys and girls). The two model approach was selected for comparison because it directly accounts for age*gender interaction that would otherwise be cumbersome. For site-effect, we performed a "site-wise" cross-validation analysis. More specifically, since there were 6 acquisition sites, we built a 6-fold nested cross-validation in such a way that data from each site was in its own fold. That is, for each iteration, the model was trained (and selected) using data from 5 sites and tested on the remaining site. The MAEs obtained for gender- and site-effects were then compared to the MAEs of the original prediction model (that includes both genders and all sites).

We explored the scope of applicability of our prediction model to functional measures, namely, intelligence quotient (IQ) scores. Performance and verbal intelligence quotient (PIQ, VIQ) and their interaction were regressed against the cross-validated residuals (estimated – chronological age) obtained from the prediction model. This experiment was designed to show whether the IQ scores that are age standardized would explain some part of the residual error in the age estimation. The dependent variables (IQ scores) were standardized to have zero-mean and unit variance before the regression analysis.

Additionally, we also investigated the potential implication of our prediction model on biological measures of maturity using Tanner scales. The Tanner scales used in the study were based on body hair and skin changes. Subjects were categorized into groups by Tanner scales (Scale *1* = Not started; *2* = Barely started; *3* = Definitely under way; *4* = Completed). We constructed age-matched groups between Tanner stages 1, 2 and 3, and compared the residual ages (estimated age – chronological age) between the groups. The aim of the experiment was to examine if a subject A with higher Tanner stage than a subject B was estimated to be older than the subject B given that the difference in the chronological age of the subjects was minimal.

Results:

Correlations and mean absolute errors (MAE) between the chronological and estimated age based on 10-fold cross-validation were computed for different resolutions with number of parcels, p = 78, 160, 640, 1248, 2560 and 10240. Across different spatial resolutions, our prediction model based on cortical thickness, produced highly accurate estimation of the chronological age. Moreover, the estimation improved with increasing spatial resolution:

correlation coefficient between the chronological and estimated age, R = 0.78, 0.81, 0.82, 0.83, 0.84 for spatial resolutions with number of cortical parcels, p = 78, 160, 640, 1248, 2560 and 10240 respectively (**Figure 1**). As seen in **Figure 1**, the improvement in estimation is also observed in the decrease of variance with increasing spatial resolutions, (variance, *var* = 0.39, 0.35, 0.32, 0.30, 0.29 for cortical parcels, p = 78, 160, 640, 1248, 2560 and 10240 respectively).

Insert **Figure 1**

The MAEs for different spatial resolutions with cortical parcels, p = 78, 160, 640, 2560 and 10240 are (in years) 1.95, 1.79, 1.74, 1.71 and 1.68 respectively. As seen in **Figure 2A**, statistical tests confirmed that the difference in the MAE was significant at the alpha level of 0.05 in all cases except between 160 and 640, 640 and 2560, 2560 and 10240 resolutions. Overall, the MAEs decreased with increasing spatial resolution. Correspondingly, a general pattern of increase in correlation between chronological and estimated age with increasing spatial resolution was observed (**Figure 2B**).

Insert **Figure 2**

The top predictors of chronological age based on cortical thickness for spatial resolutions with cortical parcels, p = 78, 160, 640, 2560 and 10240 are shown in **Figure 3A-E** and in **Table 2**. Across different spatial resolutions, we observed a largely consistent set of cortical regions in sensorimotor (precentral and postcentral gyrus), and association cortex (including middle and superior frontal gyrus) as the top predictors. It may be noted that the drop in the absolute coefficient values across the resolution level is explained by the fact that when there were more cortical parcels, the number of cortical parcels that contributed to the age estimation, was also

higher. The number of parcels with non-zero coefficient b_i in Eq. (1) were 48, 89, 106, 151, 200 for cortical atlas with p = 78, 160, 640, 2560, and 10240 respectively.

Insert **Figure 3**

Insert **Table 2**

Our exploration of the gender effect on the performance of the prediction model revealed that gender was not an important factor of the residual error between the estimated and chronological age. The "two models" approach performed significantly worse than the combined prediction-model (see **Table 3** and **supplementary Table S3**) in cortical atlas with parcels, p = 78, 160, 2560 and 10240. The performance of the site-based approach was worse than that of the combined/original prediction-model in cortical atlas with parcels, p = 78, 2560 and 10240, but the differences were not large (although they were significant).

Insert **Table 3**

As seen in **Table 4**, which shows the results of the regression analysis between IQ scores and the residuals in the age estimation for the case of 10240 cortical parcels, the IQ scores explained a small but significant part of the residual error. The corresponding results for other resolutions are relegated to the **supplementary Table S3** and they were not qualitatively different from the case of 10240 parcels. The results are not surprising since we did not expect the age estimation to be driven by the difference in intelligence, but still we expected to have a relation between the age residual and IQ scores. The regression coefficient for the interaction effect was positive indicating positive relationship between the age estimation residual and the interaction effect

(i.e., the larger the estimated age compared to the chronological age, the larger the PIQ*VIQ interaction).

Insert **Table 4**

Table 5 shows the residual age (estimated age – chronological age) for the spatial scale of 10240 parcels between Tanner scales 1, 2 and 3 that have been matched for chronological age. A significant positive difference between Tanner scales 2 and 3 (for Skin Changes) was observed. We also observed a positive difference, though non-significant, in the residual ages for increasing Tanner scales of 1, 2 and 3 for Body Hair.

Insert **Table 5**

Discussion:

In this study on a large number of normally developing children and adolescents, we used cortical thickness measured across several cortical regions to build a predictive model for estimation of chronological age via the elastic-net regularized regression that jointly performs variable selection and model estimation. Our model produced a high accuracy in the estimation of chronological age, and the accuracy improved with increasing spatial resolution as evident from increasing correlation and decreasing variance between the chronological and estimated age. Analysis of correlation and MAE between the chronological and estimated age based on 10-fold cross-validation, showed the best results at spatial resolutions consisting of 2560 and 10240 cortical parcels. Lastly, the top predictors of chronological age were observed in highly localized cortical regions including sensorimotor and association areas.

Several studies have shown cortical thickness (and/or GM density) as a sensitive index of brain development. In a landmark paper, Giedd et al. (1999), used longitudinal scans of a large number of normally developing children and adolescents to define developmental trajectories of GM volume. Subsequent studies have also shown similar developmental trajectories using cortical thickness (Gogtay et al., 2004; Shaw et al., 2008). Interestingly, trajectories of cortical anatomy have been associated with cognitive development (Gogtay et al., 2004; Shaw et al., 2006), and a delay in the developmental trajectories has been implicated in developmental disorders such as ADHD (Paus et al., 2008; Shaw et al., 2007, 2010; Giedd and Rapoport, 2010). Taken together, cortical thickness may, therefore, provide important information about maturational changes in brain; and as such, a prediction model of brain maturity based on cortical thickness may be expected to yield good estimation.

Understandably, few studies have started to use cortical thickness (Brown et al., 2012) and GM density (using voxel-based morphometry, VBM) (Erus et al., 2014; Franke et al., 2012) in prediction models of brain maturity. In the cortical thickness-based study, Brown et al. used cortical thickness averaged over the entire brain (and entire hemisphere), and as such, ascertaining which cortical regions were the top predictors of the model, was not possible. The two other studies based on GM (Erus et al., 2014; Franke et al., 2012) also did not compute the top predictors of brain maturity. The VBM approach of Franke et al., (2012) considered sub-cortical structures in addition to cortical structures in the prediction model. Also, as discussed in Hutton et al. (2009), cortical VBM provides a mixture of cortical thickness and cortical folding. Thus, the results of top predictors were less specific. Additionally, the previous studies used kernelized approach in building prediction models which made the interpretation of the resulting model challenging. The gains of assessing structural maturity based on cortical thickness and

spatially sparse predictive models as in this paper lie in the specificity of the resulting models, and clear and quantitative anatomical interpretation.

The knowledge of top predictors can reveal important insights into brain development. For example, in a recent study, Dosenbach et al. used resting state functional connectivity MRI and multi-variate pattern analysis to derive individual subject's brain maturity across development (Dosenbach et al., 2010). More importantly, the authors demonstrated that regions belonging to the cingulo-opercular control network were the best predictors of functional brain maturity. Specifically, right anterior prefrontal cortex, considered essential for cognitive control and higher-order reasoning, was found to have the largest relative prediction power. Based on the findings from functional brain imaging studies and the fact that cortical thickness is a sensitive index of brain development, we hypothesized that the top predictors of our model would reveal cortical regions that play critical role during brain maturation.

Our findings of top predictors in the primary sensorimotor and association regions are consistent with several studies that have shown developmental changes in these regions and their associated networks. Findings from human and non-human primates have shown a hierarchical sequence of structural and functional developments from sensory to motor and lastly to association cortex (Elston et al., 2009; Huttenlocher and Dabholkar, 1997; Travis et al., 2005). Imaging studies also provide evidence in support of the hierarchical sequence of cortical maturation (Gogtay et al., 2004; Lebel and Beaulieu, 2011). Additional support for the hypothesis has come from recent network-based studies that have also shown major developmental changes in large-scale structural and functional brain networks (Khundrakpam et al., 2013; Supekar et al., 2008; Zielinski et al., 2010). Zielinski et al., using structural covariance network approach, showed that primary sensorimotor connectivity were already developed in

early childhood but expanded in early adolescence before pruning to a more restricted topology resembling adult intrinsic connectivity network patterns (Zielinski et al., 2010). They also showed increasingly distributed topology of connectivity of association areas with development. In another study, Khundrakpam et al. demonstrated decreasing sensorimotor connectivity along with increasing association connectivity with development (Khundrakpam et al., 2013). Our analyses also showed significant changes in cortical thickness with age in these regions (see **Supplementary Table S4**). However, it may be noted that, while most which were highly predictive for age would be detected also in the conventional massively univariate analysis (as in the case of sensorimotor and association areas), we also identified cortical parcels that are predictive of age only in the multivariate setting, i.e., when analyzed together with other, predictive parcels (see **Supplementary Table S4**). While further variable association analysis is beyond the scope of this work, it may be noted that the complementary aspects of univariate and multivariate pattern analysis have been reported, e.g., for fMRI (Jimura and Poldrack, 2012).

We also observed that increasing spatial resolution resulted in better estimation of brain maturity. Such a finding is not surprising given recent evidence from network studies of increased topological attributes (e.g. small-worldness, efficiency) with increased spatial resolution (Tohka et al., 2012; Zalesky et al., 2010). In a recent study, Zalesky et al. (2010) defined network nodes as grey matter regions and the white matter tracts interconnecting the regions as edges; and then explore the effect of spatial resolution on several topological attributes over a wide range of nodal resolutions (number of nodes = 82, 500, 1000, 2000, 3000, 4000). The topological parameters varied considerably with spatial resolution, particularly small-worldness was found to increase by 95 % in the 4000-node scale compared to that of 82-node scale. Small-worldness, defined as the optimal topological organization of a network

characterized by high clustering coefficient and low path length (Watts and Strogatz, 1998), has been shown in structural and functional brain networks (for a detail review, see Bullmore and Sporns, 2009). The increase in small-worldness with increased spatial scale indicates that the use of higher spatial scale may reveal better organizational properties of brain architecture.

Our findings also indicate that, among the spatial scales given by AAL and CIVET's tessellation, the best estimation of brain maturity was obtained for spatial scales consisting of i) 2560 and ii) 10240 parcels. The question of spatial scale (the choice of brain parcellation) has been a topic of continuing discussion with no clear answer. Broadly, three resolutions of brain divisions on the basis of structure can be understood: microscale (level of single neurons and synapses), mesoscale (level of neuronal groups) and macroscale (level of anatomically distinct regions) (Sporns et al., 2005). The microscale consists of numerous and highly variable individual neurons and synapses, while the macroscale comprises large brain regions recognizable from the gross anatomy. Neither scale provides a sufficient basis for the understanding of the functional repertoire of human brain. In contrast, cortical minicolumns (at mesoscale) may serve as the basic functional unit of brain (Mountcastle, 1997). Such a scale of spatial resolution is likely to support *functional integration* and *functional segregation*, the two main organizing principles in human brain (Friston, 2002; Tononi et al., 1998). Taken together, among the spatial scales obtained from AAL and CIVET mesh, the spatial scales consisting of 2560 and 10240 cortical parcels are closest to mesoscale, and thus, our findings of best estimation of brain maturity at these spatial scales may indicate a level of neuronal architecture that may reflect the functional division of cortical structures.

Interestingly, adding gender information to our models seemed not to improve age prediction. We hypothesize that this occurs because it is more important to have more subjects to train the

age prediction model than to try to account for gender specific differences in the development. The result should not be interpreted in a way that the effect of gender would be insignificant for age prediction based on cortical thickness, but rather it suggests that the effect is complex and challenging to account for.

The correlation and MAE measures decreased little (though significant) when the age estimation was evaluated across sites compared to the case when the site information was ignored in the model construction and evaluation. This is consistent with an experiment presented by Franke et al. (2012) for the age estimation using voxel-based morphometry. The correlation between estimated and chronological age dropped less than two percent when cross-validation of the prediction model was performed "site-wise", indicating the strength of our prediction model in using cross-site data. This suggests that the age estimation method was robust against possible site differences demonstrating the potential to apply the method for future large-scale, multi-center studies.

In order to explore the full relevance and application of our prediction model, we explored additional measures of cortical anatomy namely surface area (SA) and cortical volume (CVol) for predicting chronological age (see **Supplementary Table S2**). The performance of the prediction model based on SA and CVol are much lower than that of CT, and so we have not included them in the main text. We believe such a difference in accuracy of prediction model (of age) is plausible given emerging evidence from recent findings that suggest that CT and SA may result from different phylogenetic processes (Pontious et al., 2008; Rakic, 1995), and their etiology may involve different genetic mechanisms (Chen et al., 2011; Chouinard-Decorte et al. 2014; Panizzon et al., 2009). Thus, it is not surprising that unique developmental trajectories (Wierenga et al., 2014) and networks (Sanabria-Diaz et al., 2010) of cortical thickness and

surface area have been shown. Taken together, CT and SA may reflect the end result of distinct phylogenetic processes which may attribute to the difference in the performance of prediction model based on CT and SA. It may however be noted that, the poor performance of the prediction model based on SA (and consequently CVol, since CVol is loosely the product of CT and SA) may be partly due to the fact that the surface area is much more difficult to measure than the cortical thickness.

Additionally, we asked whether our estimation says more about biological age than chronological age does. Our analyses revealed that residuals (estimated – chronological age) of the prediction model were significantly related to the IQ scores of the subjects although, as one might expect, the effects were not large. Our findings indicate that subjects with both high PIQ and high VIQ may display greater estimated ages than the chronological ages. These findings may thus indicate that subjects with high IQ (PIQ and VIQ) would exhibit greater brain maturity, a speculation that is not entirely implausible. In fact, in consistent with our findings, Erus et al. (2014) using T1-based regional volumetric maps of GM, WM and lateral ventricle predicted brain maturity index (BDI) in subjects (with age ranging from 8 to 22 years), and found that subjects with higher BDI than their chronological age exhibited greater cognitive processing speed (Erus et al., 2014). It may be noted that regression analyses with i) full scale intelligence quotient (FSIQ) and ii) standardized scores of WISC subtests (Block Design, Matrix Reasoning, Similarities, Vocabulary) were non-significant, and have not been included in the manuscript. We understand that the interaction term for PIQ and VIQ captures disagreement (or similarity) between the two measures (PIQ and VIQ) in relation to the residuals, which in turn might complicate the interpretation of the regression analysis as the main effects become conditional

effects. Nevertheless, we would like to point out that conditional effects of PIQ and VIQ have been studied in relation to cortical thickness, e.g., in Margolis et al. (2013).

We also demonstrated that the biological measures specifically, Tanner stages could be accurately predicted based on the estimated ages. Our findings showed that the subjects with more advanced Tanner scales were estimated to be older than age-matched coequals with lower Tanner scale.

In conclusion, our study adds to the use of MRI in studying brain developmental trajectories by providing an amalgamated index of brain maturity based on cortical thickness. The estimated ages from our prediction model relate to functional and behavioural measures, underscoring the relevance and scope of the study in the understanding of biological maturity.

Acknowledgements:

Funding: This research has been supported by The Azrieli Neurodevelopmental Research Program in partnership with Brain Canada Multi-Investigator Research Initiative (MIRI) grant to BSK and ACE, and by the Academy of Finland under the grants 130275, 263785 to JT. BSK was supported by a Post-Doctoral Fellowship from FRSQ and Jeanne-Timmins Costello MNI Fellowship. We thank the anonymous reviewers whose comments and suggestions have improved the paper.

This project has been funded in whole or in part with Federal funds from the National Institute of Child Health and Human Development, the National Institute on Drug Abuse, the National Institute of Mental Health, and the National Institute of Neurological Disorders and Stroke

(Contract #s N01-HD02-3343, N01-MH9-0002, and N01-NS-9-2314, -2315, -2316, -2317, -2319 and -2320). Special thanks to the NIH contracting officers for their support.

We also acknowledge the important contribution and remarkable spirit of John Haselgrove, Ph.D. (deceased).

Disclaimer

This manuscript reflects the views of the authors and may not reflect the opinions or views of all Study Investigators or the NIH.

Appendix A. Brain Development Cooperative Group

Key personnel from the six pediatric study centers are as follows: **Children's Hospital Medical Center of Cincinnati**, Principal Investigator William S. Ball, M.D., Investigators Anna Weber Byars, Ph.D., Mark Schapiro, M.D., Wendy Bommer, R.N., April Carr, B.S., April German, B.A., Scott Dunn, R.T.; **Children's Hospital Boston**, Principal Investigator Michael J. Rivkin, M.D., Investigators Deborah Waber, Ph.D., Robert Mulkern, Ph.D., Sridhar Vajapeyam, Ph.D., Abigail Chiverton, B.A., Peter Davis, B.S., Julie Koo, B.S., Jacki Marmor, M.A., Christine Mrakotsky, Ph.D., M.A., Richard Robertson, M.D., Gloria McAnulty, Ph.D; **University of Texas Health Science Center at Houston**, Principal Investigators Michael E. Brandt, Ph.D., Jack M. Fletcher, Ph.D., Larry A. Kramer, M.D., Investigators Grace Yang, M.Ed., Cara McCormack, B.S., Kathleen M. Hebert, M.A., Hilda Volero, M.D.; **Washington University in St. Louis**, Principal Investigators Kelly Botteron, M.D., Robert C. McKinstry, M.D., Ph.D.,

Investigators William Warren, Tomoyuki Nishino, M.S., C. Robert Almli, Ph.D., Richard Todd, Ph.D., M.D., John Constantino, M.D.; **University of California Los Angeles**, Principal Investigator James T. McCracken, M.D., Investigators Jennifer Levitt, M.D., Jeffrey Alger, Ph.D., Joseph O'Neil, Ph.D., Arthur Toga, Ph.D., Robert Asarnow, Ph.D., David Fadale, B.A., Laura Heinichen, B.A., Cedric Ireland B.A.; **Children's Hospital of Philadelphia**, Principal Investigators Dah-Jyuu Wang, Ph.D. and Edward Moss, Ph.D., Investigators Robert A. Zimmerman, M.D., and Research Staff Brooke Bintliff, B.S., Ruth Bradford, Janice Newman, M.B.A. The Principal Investigator of the data coordinating center at **McGill University** is Alan C. Evans, Ph.D., Investigators Rozalia Arnaoutelis, B.S., G. Bruce Pike, Ph.D., D. Louis Collins, Ph.D., Gabriel Leonard, Ph.D., Tomas Paus, M.D., Alex Zijdenbos, Ph.D., and Research Staff Samir Das, B.S., Vladimir Fonov, Ph.D., Luke Fu, B.S., Jonathan Harlap, Ilana Leppert, B.E., Denise Milovan, M.A., Dario Vins, B.C.,, and at **Georgetown University**, Thomas Zeffiro, M.D., Ph.D. and John Van Meter, Ph.D. Ph.D. Investigators at the Neurostatistics Laboratory, **Harvard University/McLean Hospital,** Nicholas Lange, Sc.D., and Michael P. Froimowitz, M.S., work with data coordinating center staff and all other team members on biostatistical study design and data analyses. The Principal Investigator of the Clinical Coordinating Center at **Washington University** is Kelly Botteron, M.D., Investigators C. Robert Almli Ph.D., Cheryl Rainey, B.S., Stan Henderson M.S., Tomoyuki Nishino, M.S., William Warren, Jennifer L. Edwards M.SW., Diane Dubois R.N., Karla Smith, Tish Singer and Aaron A. Wilber, M.S. The Principal Investigator of the Diffusion Tensor Processing Center at the **National Institutes of Health** is **Carlo Pierpaoli, MD, Ph.D.,** Investigators **Peter J. Basser, Ph.D., Lin-Ching Chang, Sc.D.,** Chen Guan Koay**, Ph.D. and Lindsay Walker, M.S.** The Principal Collaborators at the **National Institutes of Health** are Lisa Freund, Ph.D. (NICHD), Judith

Rumsey, Ph.D. (NIMH), Lauren Baskir, Ph.D. (NIMH), Laurence Stanford, PhD. (NIDA), Karen Sirocco, Ph.D. (NIDA) and from NINDS, Katrina Gwinn-Hardy, M.D., and Giovanna Spinella, M.D. The Principal Investigator of the Spectroscopy Processing Center at the **University of California Los Angeles** is **James T. McCracken, M.D., Investigators** Jeffry R. Alger, Ph.D., Jennifer Levitt, M.D., Joseph O'Neill, Ph.D.

References:

Ambroise, C., McLachlan, G.J., 2002. Selection bias in gene extraction on the basis of microarray gene-expression data. Proc. Natl. Acad. Sci. U.S.A. 99(10), 6562-6566.

Ameis, S.H., Ducharme, S., Albaugh, M.D., Hudziak, J.J., Botteron, K.N., Lepage, C., Zhao, L., Khundrakpam, B., Collins, D.L., Lerch, J.P., Wheeler, A., Schachar, R. Evans, A.C., Karama, S., 2014. Cortical thickness, cortico-amygdalar networks, and externalizing behaviors in healthy children. Biol. Psychiatry 75(1), 65-72.

Brown, T.T., Kuperman, J.M., Chung, Y., Erhart, M., McCabe, C., Hagler, D.J., Jr., Venkatraman, V.K., Akshoomoff, N., Amaral, D.G., Bloss, C.S., Casey, B.J., Chang, L., Ernst, T.M., Frazier, J.A., Gruen, J.R., Kaufmann, W.E., Kenet, T., Kennedy, D.N., Murray, S.S., Sowell, E.R., Jernigan, T.L., Dale, A.M., 2012. Neuroanatomical assessment of biological maturity. Curr. Biol. 22(18), 1693-1698.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10(3), 186-198.

Carroll, M.K., Cecchi, G.A., Rish, I., Garg, R., Rao, A.R., 2009. Prediction and interpretation of distributed neural activity with sparse models. NeuroImage 44(1), 112-122.

Chen, C.H., Panizzon, M.S., Eyler, L.T., Jernigan, T.L., Thompson, W., Fennema-Notestine, C., Jak, A.J., Neale, M.C., Franz, C.E., Hamza, S., Lyons, M.J., Grant, M.D., Fischl, B., Seidman, L.J., Tsuang, M.T., Kremen, W.S., Dale, A.M., 2011. Genetic influences on cortical regionalization in the human brain. Neuron 72(4), 537-544.

Chouinard-Decorte, F., McKay, D.R., Reid, A., Khundrakpam, B., Zhao, L., Karama, S., Rioux, P., Sprooten, E., Knowles, E., Kent, J.W., Jr., Curran, J.E., Goring, H.H., Dyer, T.D., Olvera, R.L., Kochunov, P., Duggirala, R., Fox, P.T., Almasy, L., Blangero, J., Bellec, P., Evans, A.C., Glahn, D.C., 2014. Heritable changes in regional cortical thickness with age. Brain Imaging Behav. 8(2), 208-216.

Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J. Comput. Assist. Tomogr. 18(2), 192-205.

Dosenbach, N.U., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., Nelson, S.M., Wig, G.S., Vogel, A.C., Lessov-Schlaggar, C.N., Barnes, K.A., Dubis, J.W., Feczko, E., Coalson, R.S., Pruett, J.R., Jr., Barch, D.M., Petersen, S.E., Schlaggar, B.L., 2010. Prediction of individual brain maturity using fMRI. Science 329(5997), 1358-1361.

Evans, A.C., Brain Development Cooperative, G., 2006. The NIH MRI study of normal brain development. Neuroimage 30(1), 184-202.

Elston, G.N., Oga, T., Fujita, I., 2009. Spinogenesis and pruning scales across functional hierarchies. J. Neurosci. 29, 3271-3275.

Erus, G., Battapady, H., Satterthwaite, T.D., Hakonarson, H., Gur, R.E., Davatzikos, C., Gur, R.C., 2014. Imaging patterns of brain development and their relationship to cognition. Cereb. Cortex. doi:10.1093/cercor/bht425

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.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. 33(1), 1-22.

Friston, K., 2002. Beyond phrenology: what can neuroimaging tell us about distributed circuitry? Annu. Rev. Neurosci. 25, 221-250.

Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., Rapoport, J.L., 1999. Brain development during childhood and adolescence: a longitudinal MRI study. Nat. Neurosci. 2(10), 861-863.

Giedd, J.N., Rapoport, J.L., 2010. Structural MRI of pediatric brain development: what have we learned and where are we going? Neuron 67(5), 728-734.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., Thompson, P.M., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. Proc. Natl. Acad. Sci. U.S.A. 101(21), 8174-8179.

Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1), 29-36.

Huttenlocher, P.R., Dabholkar, A.S., 1997. Regional differences in synaptogenesis in human cerebral cortex. J. Comp. Neurol. 387(2), 167-178.

Hutton, C., Draganski, B., Ashburner, J., Weiskopf, N., 2009. A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging. Neuroimage 48(2), 371-380.

Huttunen, H., Manninen, T., Tohka, J., 2012. MEG mind reading: Strategies for feature selection. Proceedings of the Federated Computer Science Event, 42-49.

Jimura, K., Poldrack, R.A., 2012. Analyses of regional-average activation and multivoxel pattern information tell complementary stories. Neuropsychologia 50, 544-552.

Kabani, N., Le Goualher, G., MacDonald, D., Evans, A.C., 2001. Measurement of cortical thickness using an automated 3-D algorithm: a validation study. Neuroimage 13(2), 375-380.

Khundrakpam, B.S., Reid, A., Brauer, J., Carbonell, F., Lewis, J., Ameis, S., Karama, S., Lee, J., Chen, Z., Evans, A.C., Brain Development Cooperative Group, 2013. Developmental changes in organization of structural brain networks. Cereb. Cortex 23(9), 2072-2085.

Kim, J.S., Singh, V., Lee, J.K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J.M., Kim, S.I., Evans, A.C., 2005. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. Neuroimage 27(1), 210-221.

Lebel, C., Beaulieu, C., 2011. Longitudinal development of human brain wiring continues from childhood into adulthood. J. Neurosci. 31(30), 10937-10947.

Lee, J.K., Lee, J.M., Kim, J.S., Kim, I.Y., Evans, A.C., Kim, S.I., 2006. A novel quantitative cross-validation of different cortical surface reconstruction algorithms using MRI phantom. Neuroimage 31(2), 572-584.

Lerch, J.P., Evans, A.C., 2005. Cortical thickness analysis examined through power analysis and a population simulation. Neuroimage 24(1), 163-173.

Li, F., Yang, Y., Xing, E.P., 2005. From lasso regression to feature vector machine. Advances in Neural Information Processing Systems 18, 779-786.

MacDonald, D., Kabani, N., Avis, D., Evans, A.C., 2000. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. Neuroimage, 12(3), 340-356.

Mountcastle, V.B., 1997. The columnar organization of the neocortex. Brain 120 (Pt 4), 701-722.

Mwangi, B., Hasan, K.M., Soares, J.C., 2013. Prediction of individual subject's age across the human lifespan using diffusion tensor imaging: a machine learning approach. Neuroimage 75, 58-67.

Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Jernigan, T.L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M.J., Grant, M.D., Franz, C.E., Xian, H., Tsuang, M., Fischl, B., Seidman, L., Dale, A., Kremen, W.S., 2009. Distinct genetic influences on cortical surface area and cortical thickness Cereb. Cortex 19, 2728–2735.

Paus, T., Keshavan, M., Giedd, J.N., 2008. Why do many psychiatric disorders emerge during adolescence? Nat. Rev. Neurosci. 9(12), 947-957.

Pontious, A., Kowalczyk, T., Englund, C., Hevner, R.F., 2008. Role of intermediate progenitor cells in cerebral cortex development. Dev. Neurosci. 30, 24-32.

Rakic, P., 1995. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. Trends Neurosci. 18(9), 383-388.

Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G.L., Greenstein, D., Clasen, L., Gogtay, N., Giedd, J.N., 2011. How does your cortex grow? J. Neurosci. 31(19), 7174-7177.

Raznahan, A., Toro, R., Daly, E., Robertson, D., Murphy, C., Deeley, Q., Bolton, P.F., Paus, T., Murphy, D.G., 2010. Cortical anatomy in autism spectrum disorder: an in vivo MRI study on the effect of age. Cereb. Cortex 20(6), 1332-1340.

Sanabria-Diaz, G., Melie-García, L., Iturria-Medina, Y., Alemán-Gómez, Y., Hernández-González, G., Valdés-Urrutia, L., Lídice Galán, L., Valdés-Sosa, P., 2010. Surface area and cortical thickness descriptors reveal different attributes of the structural human brain networks. Neuroimage 50, 1497–1510.

Sharda, M., Khundrakpam, B.S., Evans, A.C., Singh, N.C., 2014. Disruption of structural covariance networks for language in autism is modulated by verbal ability. Brain Struct. Funct. DOI: 10.1007/s00429-014-0953-z

Shaw, P., Eckstrand, K., Sharp, W., Blumenthal, J., Lerch, J.P., Greenstein, D., Clasen, L., Evans, A., Giedd, J., Rapoport, J.L., 2007. Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. PNAS 104(49), 19649-19654.

Shaw, P., Gogtay, N., Rapoport, J., 2010. Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. Hum. Brain Mapp. 31(6), 917-925.

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., Evans, A., Rapoport, J., Giedd, J., 2006. Intellectual ability and cortical development in children and adolescents. Nature 440(7084), 676-679.

Shaw, P., Lerch, J., Greenstein, D., Sharp, W., Clasen, L., Evans, A., Giedd, J., Castellanos, F.X., Rapoport, J., 2008. Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention-deficit/hyperactivity disorder. Arch. Gen. Psychiatry 63(5), 540-549.

Shaw, P., Malek, M., Watson, B., Greenstein, D., de Rossi, P., Sharp, W., 2013. Trajectories of cerebral cortical development in childhood and adolescence and adult attention-deficit/hyperactivity disorder. Biol. Psychiatry 74(8), 599-606.

Shaw, P., Malek, M., Watson, B., Sharp, W., Evans, A., Greenstein, D., 2012. Development of cortical surface area and gyrification in attention-deficit/hyperactivity disorder. Biol. Psychiatry 72(3), 191-197.

Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans. Med. Imaging, 17(1), 87-97.

Sporns, O., Tononi, G., Kotter, R., 2005. The human connectome: a structural description of the human brain PLoS Comput. Biol. 1, 245-51.

Supekar, K., Musen, M., Menon. V., 2009. Development of large-scale functional brain networks in children. PLoS Biol. 7:e1000157.

Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain. New York: Thieme.

Tohka, J., Zijdenbos, A., Evans, A., 2004. Fast and robust parameter estimation for statistical partial volume models in brain MRI. Neuroimage 23(1), 84-97.

Tohka, J., He, Y., Evans, A.C., 2012. The impact of sampling density upon cortical network analysis: regions or points. Magn. Reson. Imaging 30(7), 978-992.

Tononi, G., Edelman, G.M., Sporns, O., 1998. Complexity and coherency: integrating information in the brain. Trends Cogn. Sci. 2(12), 474-484.

Tononi, G., Edelman, G.M., Sporns, O., 1998. Complexity and coherency: integrating information in the brain. Trends Cogn. Sci. 2(12), 474-484.

Travis, K., Ford, K., Jacobs, B., 2005. Regional dendritic variation in neonatal human cortex: a quantitative Golgi study. Dev. Neurosci. 27, 277-287.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM

using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15(1), 273-289.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393, 440-442.

Wierenga, L.M., Langen, M., Oranje, B., Durston, S., 2014. Unique developmental trajectories of cortical thickness and surface area. Neuroimage, 87, 120-126.

Zalesky, A., Fornito, A., Harding, I.H., Cocchi, L., Yücel, M., Pantelis, C., Bullmore, E.T., 2010. Whole-brain anatomical networks: does the choice of nodes matter? Neuroimage 50(3), 970-83.

Zielinski. B.A., Gennatas, E.D., Zhou, J., Seeley, W.W., 2010. Network-level structural covariance in the developing brain. PNAS 107, 18191-18196.

Zijdenbos, A.P., Forghani, R., Evans, A.C., 2002. Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. IEEE Trans. Med. Imaging, 21(10), 1280-1291.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Series B Stat. Methodol. 67, 301-320.

Figure Legends:

Fig. 1. Prediction of chronological age based on cortical thickness. ***A-E*** denote the prediction results for different spatial resolutions (number of cortical parcels = 78, 160, 640, 2560, 10240). For each figure, the chronological age is shown in x-axis, and estimated age in y-axis (upper row); while the lower row shows the residuals of the fitted model. An increase in prediction accuracy (increased '*R*' and decreased '*var*') was observed with increased spatial resolution. *R* denotes correlation coefficient and *var* denotes variance.

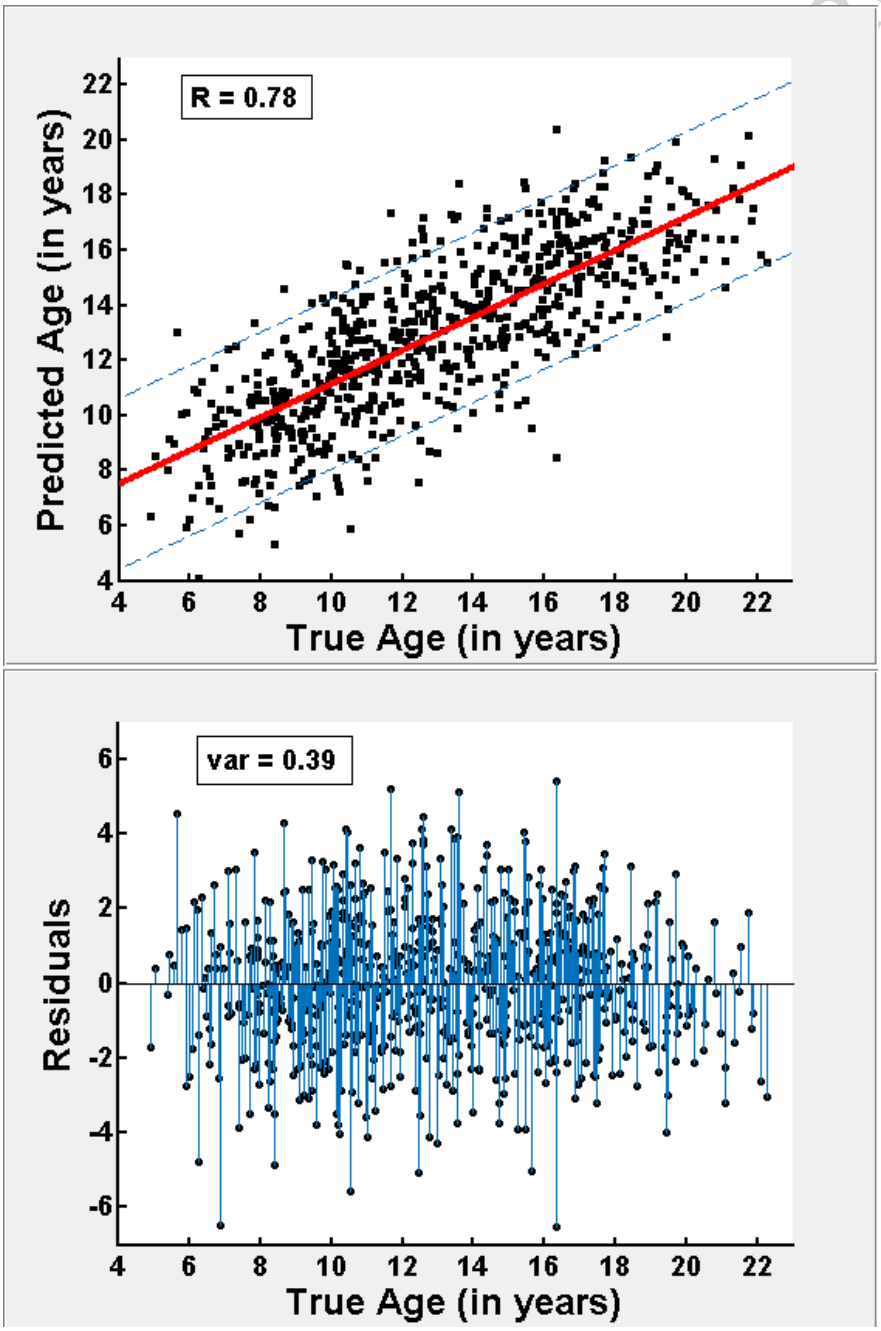
Fig. 2. Effect of spatial resolutions on the prediction model. ***A*** Mean Absolute Errors (MAEs) between the estimated and chronological age based on 10-fold cross-validation. ***B*** Correlation between the estimated and chronological age based on 10-fold cross-validation. For statistical comparisons, the reader is referred to the **Methods** section. With increasing spatial resolutions, a general pattern of decreased MAE and increased correlation was observed.

Fig. 3. Top predictors of age estimation. ***A-E*** Top predictors for different spatial resolutions (number of cortical parcels = 78, 160, 640, 2560, 10240). To obtain a single model, all subjects were used to create the model and the parameter λ was decided based on 10-fold cross-validation. Top predictors were decided based on the absolute values of the model coefficients b_i. Note that the absolute value of the coefficient b_i can be used to assess the relative importance of ROI/predictor *i* in the model with respect to other predictors because the cortical thickness measurements were standardized (see **Methods**). A largely consistent set of cortical regions in sensorimotor and association cortices were observed as the top predictors across the spatial resolutions. Note that arrows have been used in **3*E*** to highlight the highly localized predictors.

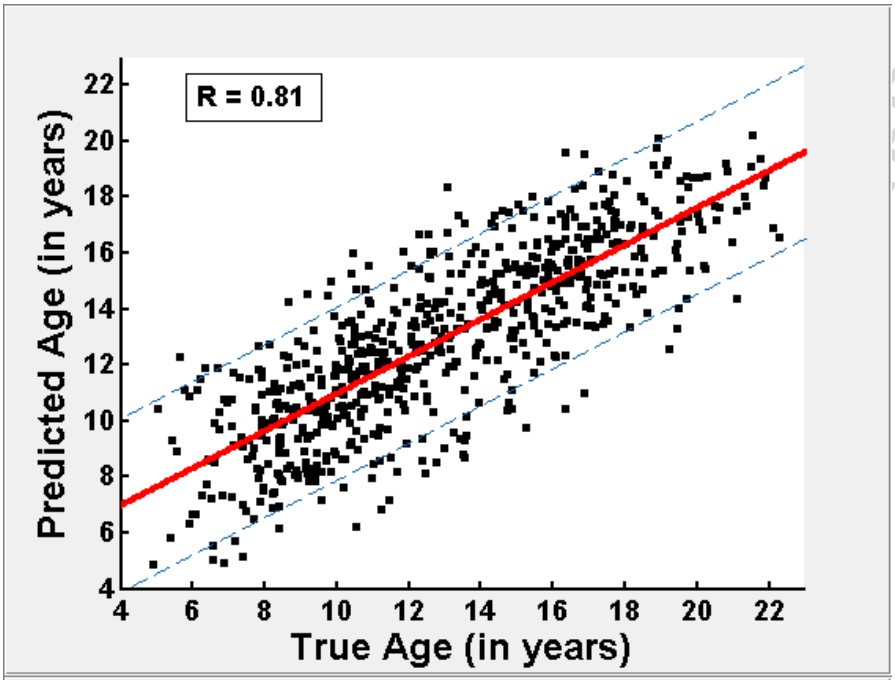
Figures:

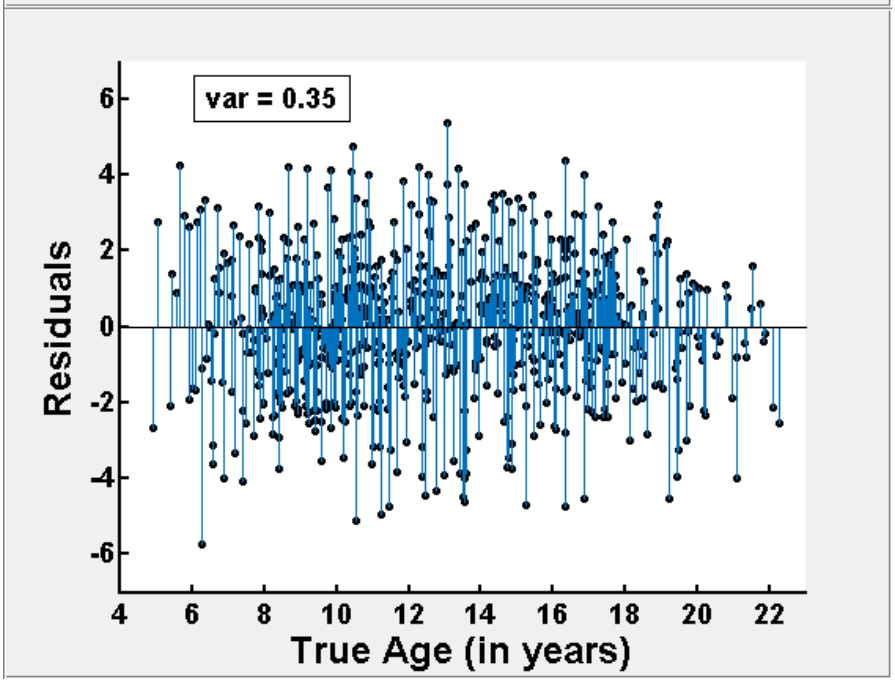
Figure 1.

(A)

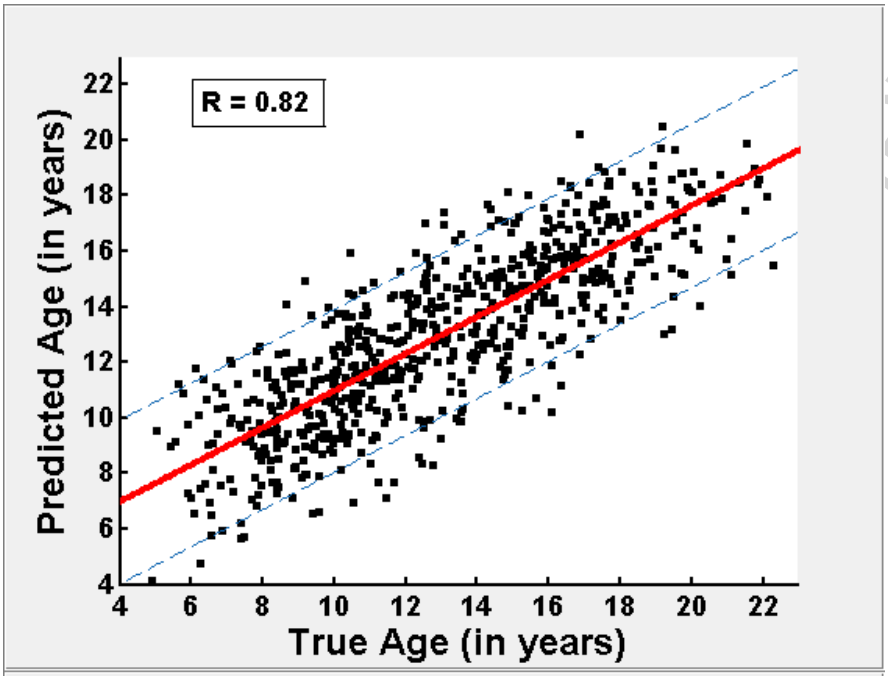


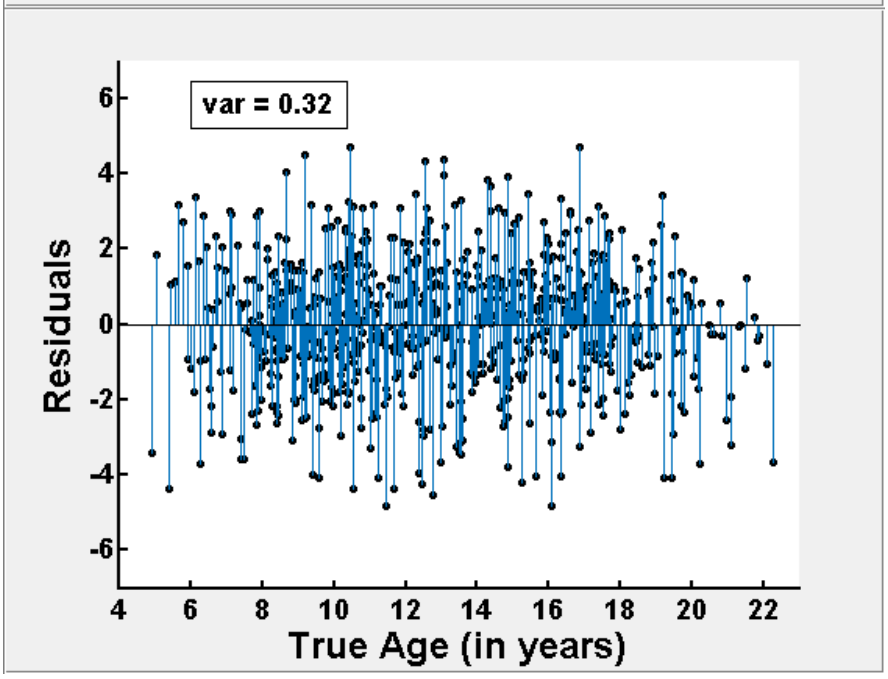
(B)



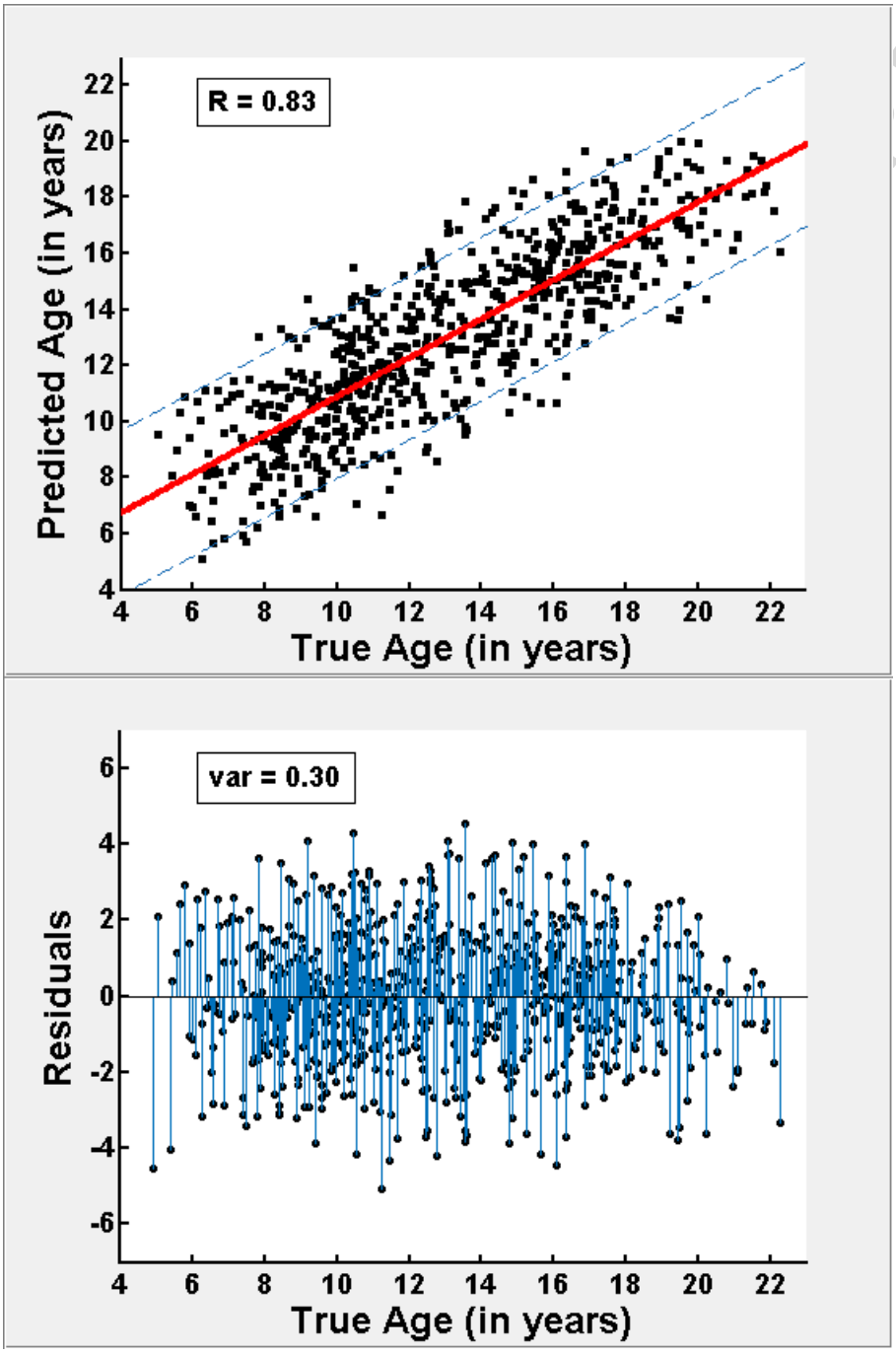


(C)





(D)



(E)

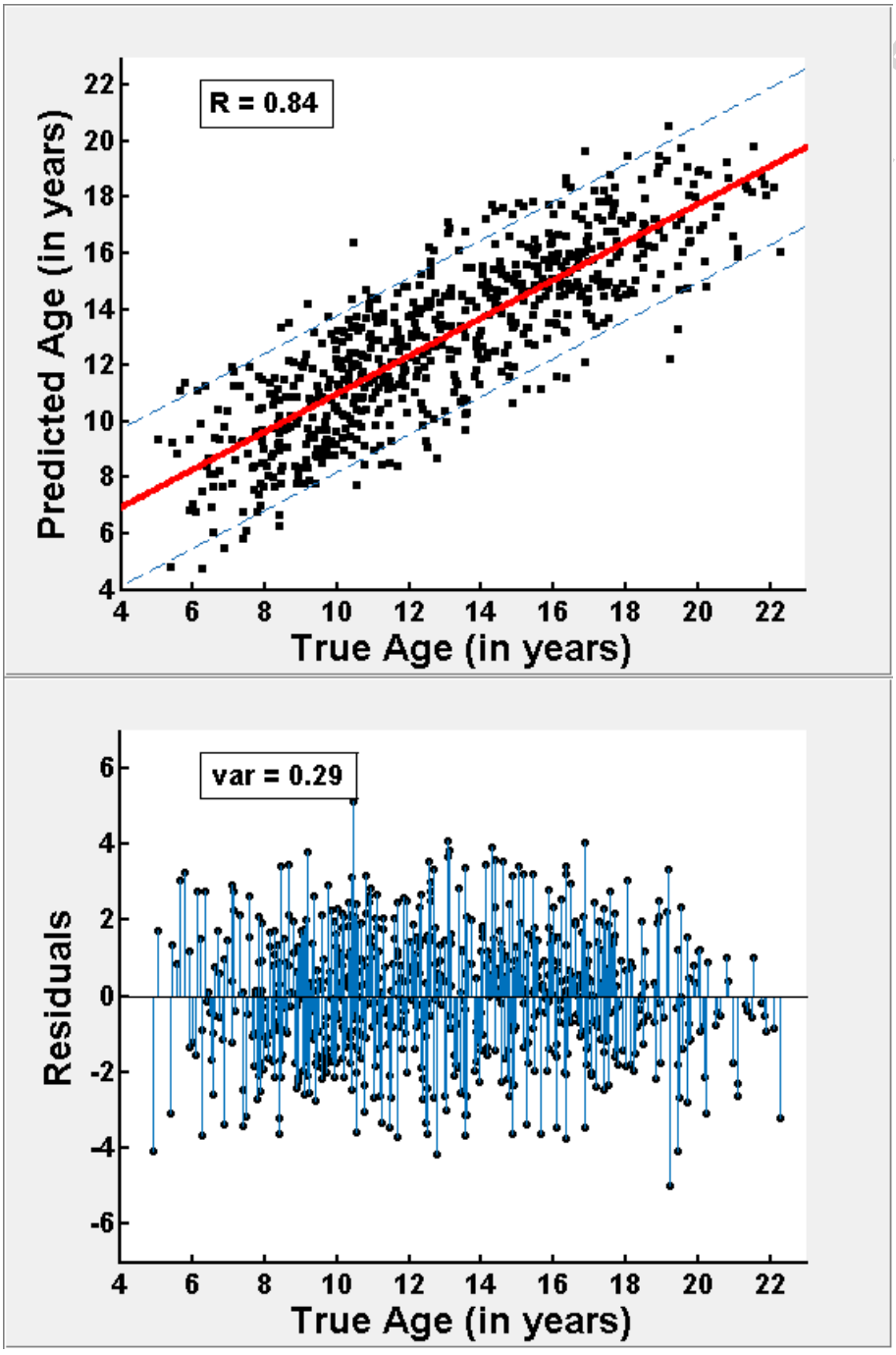
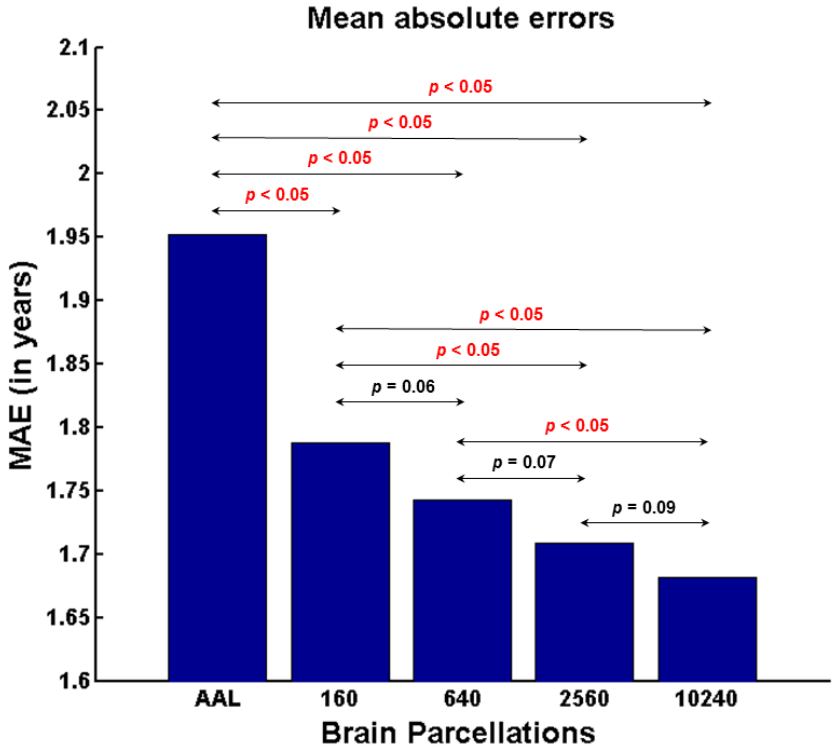


Figure 2.

(A)



(B)

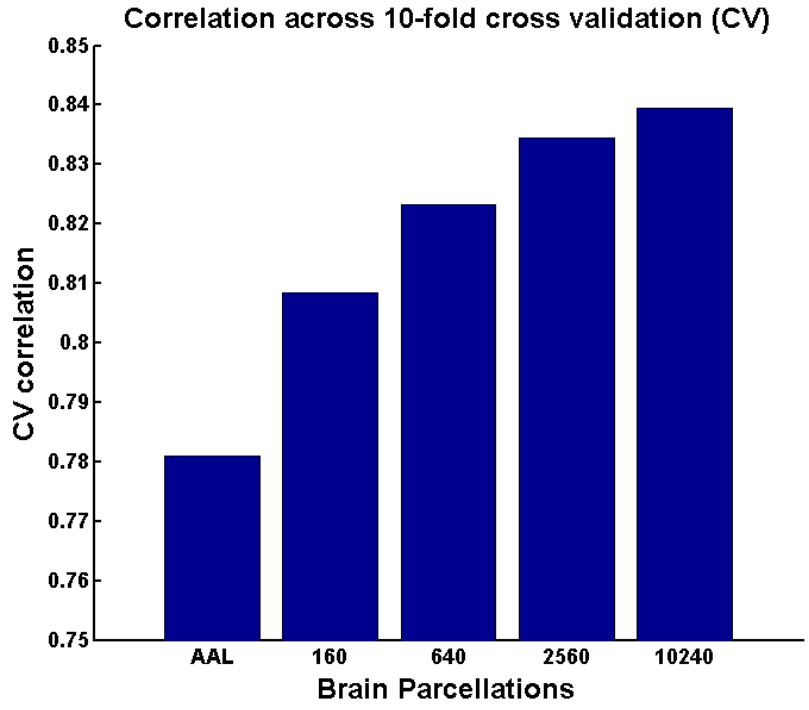
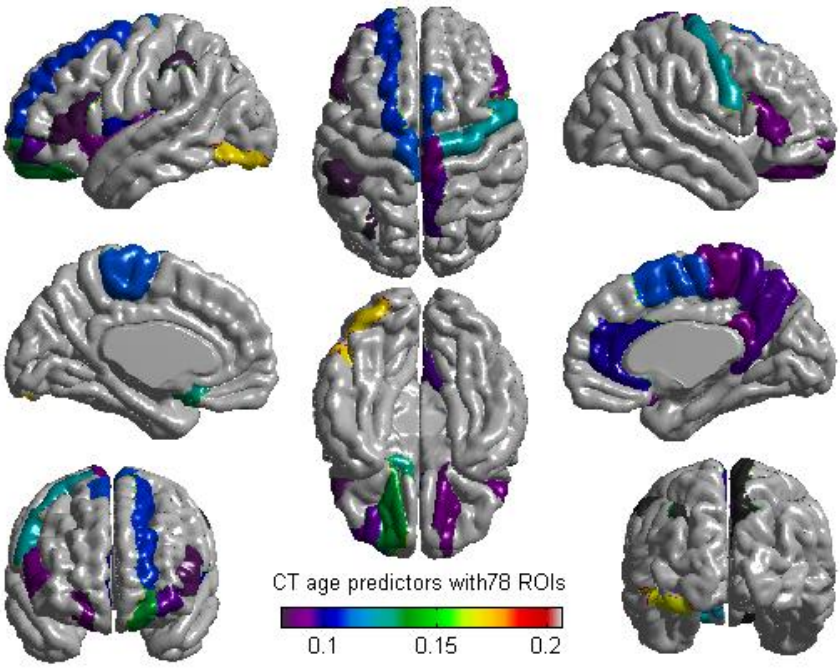
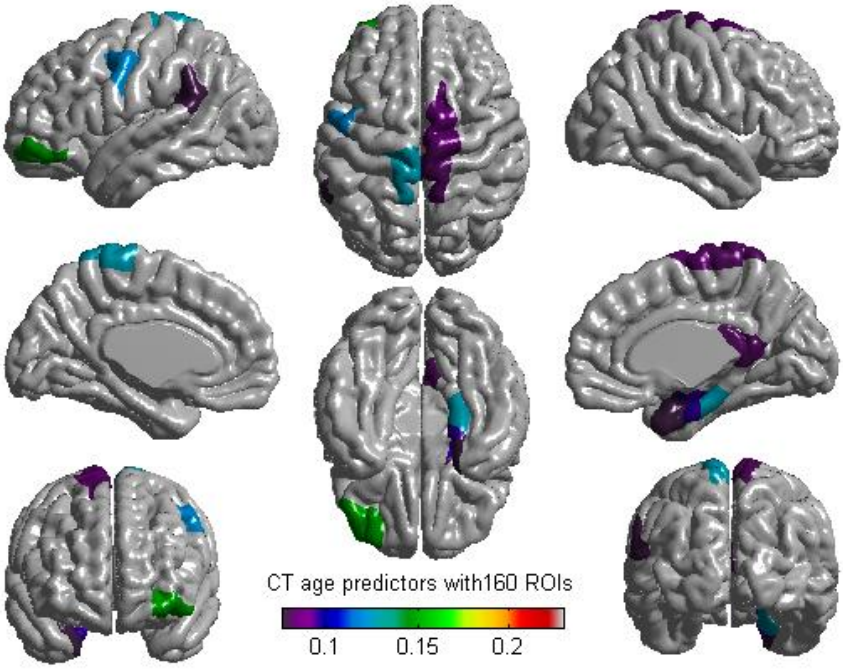


Figure 3.

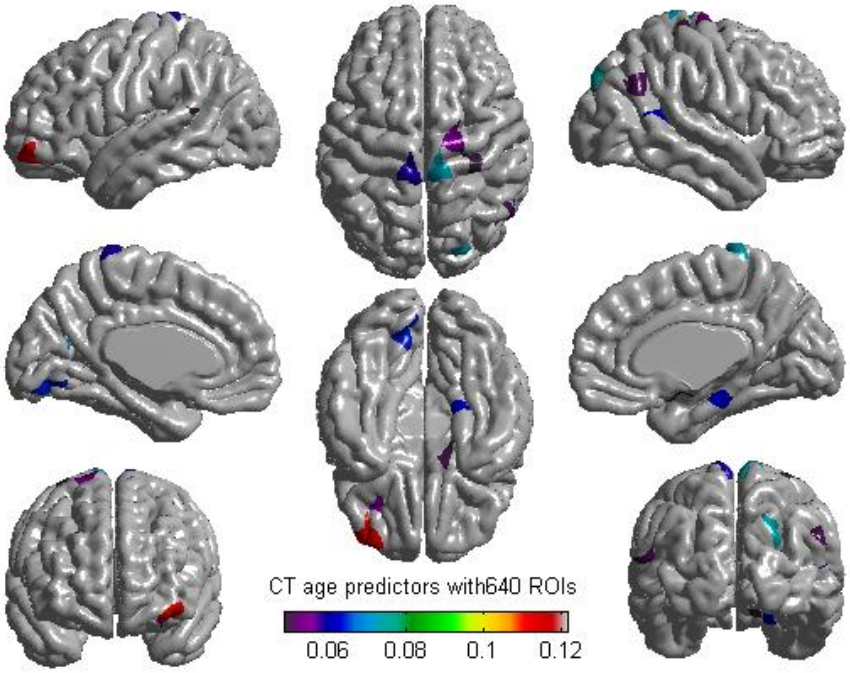
(A)



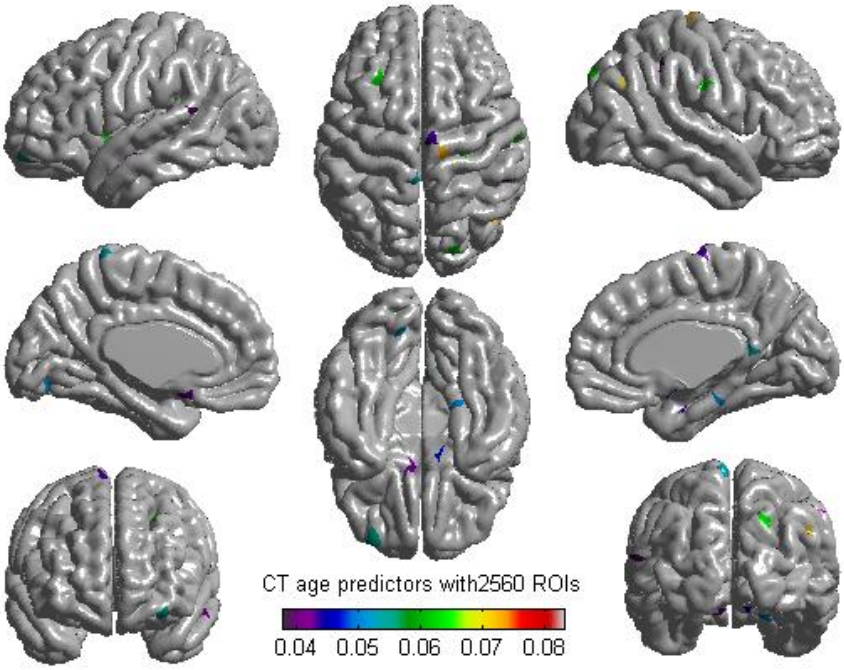
(B)



(C)



(D)



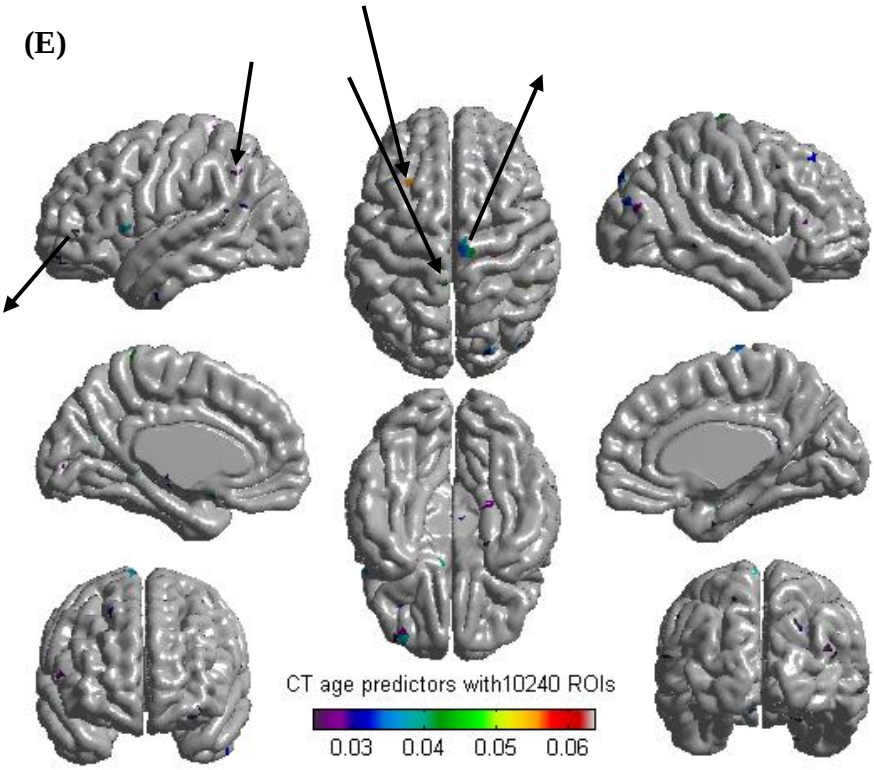


Table 1: Demographic details of subjects used in the study. FSIQ = full scale intelligence quotient, PIQ = performance intelligence quotient, VIQ = verbal intelligence quotient.

Total number of subjects (Males/Females):	308 (136/172)
Total number of scans:	679
Total number of acquisition sites:	6
Age:	12.9 ± 3.8
FSIQ:	111.7 ± 12.1
PIQ:	110.6 ± 12.7
VIQ:	110.3 ± 12.9

Table 2: Table depicting the top 10 regions of interest (ROIs) for predicting chronological age. The ranking is based on the absolute values of model coefficients b_i. To obtain a single model, all subjects were used to create the model and the parameter λ was decided based on 10-fold cross-validation (see **Methods**). Note that most of the regions are consistently the top predictors of biological age across the different spatial resolutions.

78 ROIs

Rank	Coefficient	X	Y	Z	Brain Label
1	0.20	7.1	-31.3	64.6	Right Paracentral Lobule
2	0.16	41.0	-7.4	48.4	Right Precentral Gyrus
3	0.12	-34.4	51.2	-8.4	Left Middle Frontal Gyrus
4	0.11	-9.1	-58.0	36.7	Left Precuneus
5	0.09	-45.6	-62.7	39.6	Left Angular Gyrus
6	0.09	17.5	41.6	-17.2	Right Superior Frontal Gyrus
7	0.09	-19.1	42.1	-17.1	Left Superior Frontal Gyrus
8	0.09	-24.1	-59.7	55.9	Left Superior Parietal Gyrus
9	0.08	-51.1	-29.1	-25.4	Left Inferior Temporal Gyrus
10	0.08	-43.6	-23.8	49.7	Left Postcentral Gyrus

160 ROIs

Rank	Coefficient	X	Y	Z	Brain Label
1	0.23	51.5	-7.6	42.4	Right Precentral Gyrus
2	0.14	-36.1	51.6	-12.3	Left Middle Frontal Gyrus
3	0.12	-51.9	-10.8	43.2	Left Postcentral Gyrus
4	0.09	32.9	-11.2	-21.2	Right Parahippocampal Gyrus
5	0.08	13.2	-7.9	72.4	Right Superior Frontal Gyrus
6	0.08	3.7	-46.2	12.9	Right Posterior Cingulate
7	0.07	-60.7	-46.2	26.3	Left Supramarginal Gyrus
8	0.06	17.0	-88.6	31.1	Right Cuneus
9	0.06	-38.9	10.0	-4.8	Left Insula
10	0.06	-3.2	17.6	-18.2	Left Medial Frontal Gyrus

640 ROIs

Rank	Coefficient	X	Y	Z	Brain Label
1	0.12	51.5	-7.6	42.4	Right Precentral Gyrus
2	0.11	-36.1	51.6	-12.3	Left Middle Frontal Gyrus
3	0.10	-40.7	-18.7	43.7	Left Precentral Gyrus
4	0.07	25.8	-83.0	37.2	Right Cuneus
5	0.06	-19.4	-64.5	20.6	Left Precuneus
6	0.06	48.5	-45.2	14.9	Right Superior Temporal Gyrus
7	0.05	-28.8	34.6	-11.7	Left Middle Frontal Gyrus
8	0.05	18.3	-12.5	72.2	Right Superior Frontal Gyrus
9	0.05	-53.6	-53.1	19.1	Left Superior Temporal Gyrus
10	0.05	14.4	1.6	-14.5	Right Precentral Gyrus

2560 ROIs

Rank	Coefficient	X	Y	Z	Brain Label
1	0.08	-19.4	-64.5	20.6	Left Precuneus
2	0.07	-40.7	-18.7	43.7	Left Precentral Gyrus
3	0.07	47.8	-65.3	34.3	Right Angular Gyrus
4	0.06	-40.4	2.4	-1.3	Left Insula
5	0.06	59.6	-15.4	32.6	Right Postcentral Gyrus
6	0.06	22.3	-86.4	41.1	Right Precuneus
7	0.06	26.5	-26.0	64.8	Right Precentral Gyrus
8	0.06	-27.8	21.6	48.0	Left Middle Frontal Gyrus
9	0.05	-30.6	57.4	-15.2	Left Superior Frontal Gyrus
10	0.05	3.7	-46.2	12.9	Right Posterior Cingulate

10240 ROIs

Rank	Coefficient	X	Y	Z	Brain Label
1	0.06	47.8	-65.3	34.3	Right Angular Gyrus
2	0.06	12.9	-0.2	-13.6	Right Parahippocampal Gyrus
3	0.06	-10.6	-78.3	-7.3	Left Lingual Gyrus
4	0.06	28.2	-25.6	63.0	Right Precentral Gyrus
5	0.05	-40.0	-2.0	-2.7	Left Insula
6	0.05	-27.8	21.6	48.0	Left Middle Frontal Gyrus
7	0.05	-19.4	-64.5	20.6	Left Precuneus
8	0.05	-40.4	-18.3	41.5	Left Precentral Gyrus
9	0.04	47.5	-44.6	50.1	Right Inferior Parietal Lobule
10	0.03	-40.7	-18.7	43.7	Left Precentral Gyrus

Table 3: Summary of comparative analyses for gender- and site-effects on the performance of the prediction model. Original refers to the method as in the paper (10-fold cross-validation which included both boys and girls in the model). Site refers to the case when CV folds are site-based (6-fold CV), while gender refers to the case when different models are used for boys and girls. * denotes significant difference ($p < 0.05$) in the comparison of Original - MAE vs Site – MAE and Original - MAE vs Gender - MAE.

Scale	**AAL**	**160**	**640**	**2560**	**10240**
Original - Correlation	0.7811	0.8084	0.8233	0.8344	0.8395
Site - Correlation	0.7689	0.7976	0.8186	0.8181	0.8227
Gender - Correlation	0.7592	0.7988	0.8272	0.8166	0.8218
Original - MAE (days)	713.0096	653.0294	636.6958	623.9933	614.0718
Site - MAE (days)	728.1468*	664.4780	646.6406	651.2781*	643.1782*
Site - *p*-value	0.0090	0.0654	0.0886	0.0012	0.0002
Gender - MAE (days)	732.3580*	676.8428*	632.1048	658.4153*	643.4086*
Gender - *p*-value	0.0439	0.0249	0.6438	0.0087	0.0154

Table 4: Regression of IQ and residuals of estimated age. The table shows the regression of IQ (PIQ, VIQ, PIQ*VIQ) and age residual (Estimated Age – Chronological Age) for spatial resolution of 10240 parcels. The effect of IQ is significant albeit small, higher the interaction term, higher the residual. Note, the coefficients are standardized.

Linear Regression Model:

Age Residual ~ 1 + PIQ + VIQ + PIQ*VIQ

Estimated Coefficients (10240 parcels)

	Estimate	**Standard Error**	***t*-statistics**	***p*-value**
Intercept	0.4	29.2	0.0	0.9886
PIQ	-587.7	218.1	-2.7	0.0072
VIQ	-636.6	224.9	-2.8	0.0047
PIQ*VIQ	1030.6	369.0	2.8	0.0049

Number of observations: 669, Error degrees of freedom: 665,

Root mean squared error: 756, R^2: 0.012, Adjusted R^2: 0.0076,

F-statistic vs constant model: 2.7, *p*-value: 0.0448

Table 5: Residual age (Estimated age – Chronological age) between Tanner scales 1, 2 and 3 (matched for chronological age) for spatial scale of 10240 parcels. Residual^1, Residual^2 and Residual^3 denote residual age (measured in days) for Tanner scales 1, 2 and 3 respectively. The Tanner scales were defined as *1* = Not started, *2* = Barely started, *3* = Definitely under way, and *4* = Completed.

A. Skin Changes (Tanner Scale 2 vs Tanner Scale 3)

Mean (Residual^3 – Residual^2) = 380.3163

p-value = 0.0017, *T*-statistic = 3.3257

Degrees of freedom, *df* = 49, Standard deviation, *sd* = 808.6213

B. Skin Changes (Tanner Scale 1 vs Tanner Scale 2)

Mean (Residual^2 – Residual^1) = 12.4401

p-value = 0.9244, *T*-statistic = 0.0954

Degrees of freedom, *df* = 40, Standard deviation, *sd* = 834.6913

C. Body Hair (Tanner Scale 1 vs Tanner Scale 2)

Mean (Residual^2 – Residual^1) = 184.5232

p-value = 0.2350, *T*-statistic = 1.2057

Degrees of freedom, *df* = 40, Standard deviation, *sd* = 979.9739

D. Body Hair (Tanner Scale 2 vs Tanner Scale 3)

Mean (Residual^3 – Residual^2) = 104.7658

p-value = 0.3051, *T*-statistic = 1.0332

Degrees of freedom, *df* = 70, Standard deviation, *sd* = 854.4026