NeuroImage 111 (2015) 350–359

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Prediction of brain maturity based on cortical thickness at different spatial resolutions

Budhachandra S. Khundrakpam [a,*,1], Jussi Tohka [b,1], Alan C. Evans [a], Brain Development Cooperative Group [2]

[a] Montreal Neurological Institute, McGill University, Montreal, Canada
[b] Department of Bioengineering and Aerospace Engineering, Universidad Carlos III de Madrid, Spain

ARTICLE INFO

Article history:
Accepted 19 February 2015
Available online 28 February 2015

Keywords:
Cortical thickness
Prediction model
Brain maturation
Structural magnetic resonance imaging
Elastic-net regularized regression

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 10,240 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 10,240 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.

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. (2014) 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

* Corresponding author at: Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada H3A 2B4.
E-mail address: budha@bic.mni.mcgill.ca (B.S. Khundrakpam).
[1] Equal authors.
[2] See Appendix for author list and affiliations of the Brain Development Cooperative Group.

http://dx.doi.org/10.1016/j.neuroimage.2015.02.046
1053-8119/

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 and Brain Development Cooperative, 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.

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 1 mm 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 (2 mm) 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 3 mm (Evans and Brain Development Cooperative, 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 (163,840 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.

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

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 10,240. 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 10,240 were obtained by recursively merging the neighbouring triangles of the triangular surface mesh model.

The prediction models contained up to 10,241 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 $x_i = [x_{i1}, \ldots, x_{ip}]^T$, and the model parameters by $\beta = [b_1, \ldots, b_p]^T$. We aim to minimize the elastic net cost function, which is written as (Zou and Hastie, 2005) –

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

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 $||\beta||_1 = \sum_{j=1}^{p}(|b_j|)$ 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