

**Working title:** Estimating the age of healthy adolescents from T1-weighted MRIs and machine learning methods

**Authors:** Maxwell Reynolds<sup>1</sup>, Joy Roy<sup>1</sup>, Rafael Ceschin<sup>2</sup>

<sup>1</sup> University of Pittsburgh School of Medicine, Department of Biomedical Informatics

<sup>2</sup> UPMC Children's Hospital of Pittsburgh, Department of Radiology

**Objective:** Implement a framework for automatically estimating the age of adolescents from T1 weighted MRI scans and a suite of machine learning methods.

**Motivation:**

Accurate modeling of a healthy patient's age based on brain imaging is an explored problem. Over the last 10 years, several studies have attempted to understand the normal aging process in humans and use it for early identification of pathologic brain developments. Most of this work has been dedicated to adult patients and the prediction of Alzheimer's Disease; however, there has been a lack of focus on development and age prediction specifically in adolescents. We will attempt to fill this gap and accurately predict normal adolescent ages from structural MRIs of their brains. This provides a framework for describing delayed or accelerated brain maturation, namely in patients diagnosed with Congenital Heart Disease whose brains appear to have delayed development.

**Methods:**

*Data*

The data for this study will be retrieved from the ABCD Repository which houses all the data from the Adolescent Brain and Cognitive Development (ABCD) study. The ABCD Study is a longitudinal investigation on brain development conducted by 21 different investigators from across the United States (Casey et al, 2018). The data we will be requesting from the repository are the minimally processed T1-weighted MRI scans of healthy adolescents' brains (ages 12 to 19). We will request access to the repository through the University of Pittsburgh's Office of Sponsored Research (OSR) who will help submit a Data Access Request to the National Institute of Mental Health Data Archive (NDA) who maintains the repository. Dr. Rafael Ceschin will help communicate with the OSR and will be the designated Principal Investigator of the study.

*Preprocessing*

Preprocessing of images will be done with the use of the FMRIB Software Library (FSL) and Python 3.6. All T1 weighted images will have the brains extracted, and gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) normalized and segmented within the same generative model. The bias fields will be corrected and the images will be registered with an affine registration (AF) to the MNI152 Atlas. Finally, to reduce noise and size, images will be passed through a gaussian filter and downsampled.

*Prediction*

To predict the ages, we will implement a number of machine learning tools and benchmark their effectiveness independently. This study will employ a supervised machine learning approach where we will supply the models a training set of images and their corresponding ages, and measure their effectiveness with a test set. The models includes Relevance Vector Regression

(RVR), Support Vector Regression (SVR), Linear Regression with L1 and L2 Regularization (Elasticnet), and finally, time permitting a 3Dimensional Convolutional Neural Network (3DCNN). SVRs and RVRs are both kernel based methods used in the original framework introduced by Franke et al. 2010. We also intend to use a linear regression model as this is a foundational method widely used and a good benchmark to compare the more advanced techniques. Finally, because deep learning methods have accomplished several feats in recent years, especially in the field of image classification and regression, we think it would be proper to also include a CNN to our list of methods. Since this is a regression task, the best way to compare the performances is to compute a Mean Absolute Error(MAE) for each tool. Lower errors would indicate better performance.

### *Application to CHD*

Term born neonates born with Congenital Heart Disease (CHD) have been shown to have brains that are more akin to healthy preterm neonates (Panigrahy et al. 2016). Time permitting, once we develop a tool that satisfactorily predicts age of healthy adolescent patients, we will predict the ages of adolescents with history of CHD. In this way, we can test the potential of our tool to provide clinically relevant descriptions on the effects of CHD compared to healthy adolescents.

### **Literature Review:**

Franke et al. (2010) proposed an initial relevance vector machine (RVM) model for predicting healthy adults' (ages 19-86) age from their T-1 weighted brain MRI image with a mean absolute error of 5 years. Their pipeline involved several pre-processing steps including bias field correction, spatial normalization, grey matter segmentation, affine registration, a smoothing kernel, and PCA dimensionality reduction. Their method showed similar test accuracy for new scanners (not seen in training data). Additionally, relevance vector regression (RVR), the Bayesian alternative to support vector regression, provides probabilistic outputs and requires only one parameter (the kernel function).

Franke et al. (2012) extended their method to children (ages 5-18) and showed similar accuracy at estimating age as well as scanner invariance while still using a single T1-weighted brain MRI image (MAE=1.1, r=0.93). Brown et al. (2012) used multimodal brain imaging data in children (ages 3-20) to predict age with similar accuracy (MAE=1.0, r=0.96) using a model based on quantitative metrics like cortical thickness, segmented volumes, signal intensity, and water diffusivity. T1-, T2-, and diffusion-weighted images were used.

Recent approaches have used CNNs to predict brain age. Cole et al. (2017) demonstrated that CNNs can be used to accurately predict brain age (MAE=4.65, r=.94) without any preprocessing (beyond a rigid registration to ensure consistent orientation and spline interpretation to obtain a common voxel size and dimension). Huang et al. similarly used a VGG-based CNN to obtain high accuracy in healthy adults (MAE=4.0). Jonsson et al (2019) trained 4 ResNet-based CNNs for brain age, each corresponding with a different kind of image (T1, Jacobian map, GM, WM), then trained a simple linear regression model using the outputs from each CNN to obtain a weighted average brain age for each adult patient (MAE=3.518). Hong et al. (2020) use a CNN to predict brain age in children ages 0-5 with MAE=67.6 days, r=0.985.

Brown TT, Kuperman JM, Chung Y, et al. Neuroanatomical assessment of biological maturity. *Curr Biol*. 2012;22(18):1693-1698. doi:10.1016/j.cub.2012.07.002

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., ... Dale, A. M. (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, 32, 43–54. doi:10.1016/j.dcn.2018.03.001

Cole JH, Poudel RPK, Tsagkrasoulis D, et al. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *Neuroimage*. 2017;163(March):115-124. doi:10.1016/j.neuroimage.2017.07.059

Franke, K., Ziegler, G., Klöppel, S., Gaser, C., & Alzheimer's Disease Neuroimaging Initiative (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. Brain maturation: Predicting individual BrainAGE in children and adolescents using structural MRI. *Neuroimage*. 2012;63(3):1305-1312. doi:10.1016/j.neuroimage.2012.08.001

Franke, K., & Gaser, C. (2019). Ten Years of BrainAGE as a Neuroimaging Biomarker of Brain Aging: What Insights Have We Gained?. *Frontiers in neurology*, 10, 789. <https://doi.org/10.3389/fneur.2019.00789>

Hong J, Feng Z, Wang SH, et al. Brain Age Prediction of Children Using Routine Brain MR Images via Deep Learning. *Front Neurol*. 2020;11(October):1-13. doi:10.3389/fneur.2020.584682

Huang T, Chen H, Fujimoto R, et al. AGE ESTIMATION FROM BRAIN MRI IMAGES USING DEEP LEARNING Department of Computer Science , National Tsing-Hua University , Taiwan Graduate School of Information Science , Tohoku University , Japan South China University of Technology , China Institute of D. *Conf 2017 IEEE 14th Int Symp Biomed*. 2017;2(1):849-852.

Jonsson BA, Bjornsdottir G, Thorgeirsson TE, et al. Brain age prediction using deep learning uncovers associated sequence variants. *Nat Commun*. 2019;10(1):1-10. doi:10.1038/s41467-019-13163-9

Panigrahy A, Lee V, Ceschin R, Zuccoli G, Beluk N, Khalifa O, Votava-Smith JK, DeBrunner M, Munoz R, Domnina Y, Morell V, Wearden P, Sanchez De Toledo J, Devine W, Zahid M, Lo CW. Brain Dysplasia Associated with Ciliary Dysfunction in Infants with Congenital Heart Disease. *J Pediatr*. 2016 Nov;178:141-148.e1. doi: 10.1016/j.jpeds.2016.07.041. Epub 2016 Aug 26. PMID: 27574995; PMCID: PMC5085835.

Tipping ME. The relevance vector machine. *Adv Neural Inf Process Syst*. 2000;(x):653-658.

**Dummy Results:**

**Table 1.** Data Demographics

	Training Set	Test Set
No. subjects		
Males/females		
Age mean (SD)		
Age range		

**Table 2.** Comparison of Regression Methods

	SVR	RVR	Lin Reg	CNN
Mean Absolute Error (MAE)				
Root Mean Squared Error (RMSE)				
Correlation (R)				
Confidence Interval (CI)				