Predicting BrainAGE in adolescence from metrics of brain change

Introduction

Over the last twenty years, research in adolescent brain development has been rapidly accelerated by technological and methodological advancements, particularly related to structural magnetic resonance imaging. As the field has advanced knowledge around developmental trajectories of brain development, there has been interest in creating metrics that can assess brain maturity. One proposed metric is the Brain-Age Gap Estimation (BrainAGE), which reflects the difference between someone's chronological age and their age as predicted by a machine learning model trained on neuroimaging data, often structural MRI (Brown et al., 2012).

BrainAGE was originally proposed in the aging literature, where an older-appearing brain has been associated with greater risk for Alzheimer's and cognitive decline (Franke et al., 2012). However, in recent years BrainAGE has been applied in the context of adolescent development, where alterations in developmental trajectories are hypothesized to relate to risk for psychopathology (Shaw et al., 2010). Indeed, adolescent BrainAGE has been found to relate to anxiety, depression, and obsessive-compulsive symptoms, among other outcomes (Cohen et al., 2023, Drobinin et al., 2022, Cropley et al., 2021).

Adolescent BrainAGE models work by harnessing known trajectories of developmental change to predict age from brain features. During adolescence, global cortical volume, thickness, and surface area decrease, while white matter volume increases (Mills et al., 2016). Additionally, there are a number of region-specific changes in these metrics, including in subcortical regions (Tamnes et al., 2017). As the brain undergoes age-related changes, the premise of the adolescent BrainAGE framework is that someone's age can be inferred from their brain measurements, and that a difference between predicted and chronological age may reflect alterations from the expected developmental trajectory, which could be related to health outcomes.

Notably, the adolescent BrainAGE literature has almost exclusively used cross-sectional data to train and test models. The only identified study to incorporate specific measures of change is Wierenga et al. (2019), which created a model that incorporated both cross-sectional measures and a maturational index (MI), defined as the slope of change for a measure averaged across two time windows. They found that the model that utilized both MI and cross-sectional data from the final time point performed better (as defined by a lower MAE) than a model using only data from the final time point, when predicting age at the final time point. Despite the improvement in performance, this method has not been widely applied. While it's unclear precisely why this method hasn't been adopted, a potential cause could be the barriers to using longitudinal data, such as a lack of available multi-time point datasets with sufficient sample sizes to train a BrainAGE model.

This use of only cross-sectional data presents an issue in this context. Longitudinal work has found evidence for non-linear trajectories of change, indicating steeper change at certain points

in adolescence in the measures commonly used as features in BrainAGE models, including cortical volume and surface area (Mills et al., 2016). Additionally, recent longitudinal work has found evidence for a large degree of individual variation in typical adolescent brain development, indicating that single measures may not provide the full picture of development (Mills et al., 2021). Adolescent development research has also begun to move towards a focus on longitudinal work as a whole, as it allows for examining individual differences in development, as well as trajectories of developmental change, which cannot be investigated using cross-sectional data (Becht & Mills, 2020).

In the context of BrainAGE, cross-sectional data is a strong limitation, as we are unable to establish whether an individual is on an altered trajectory of development, or is following a normative trajectory with higher or lower starting values for a given measure. For example, (non-linear) decreases in cortical volume have been observed in adolescence (Mills et al., 2016). Under the current framework, a smaller cortical volume is assumed to relate to an older chronological age, under the assumption that the individual has undergone cortical volume reduction. However, when using cross-sectional data, we can't know whether that individual has actually undergone a cortical volume reduction, or has a smaller cortical volume to start with and has not yet experienced a reduction. As BrainAGE has been interpreted as reflecting brain maturation and proposed to assess accelerated or decelerated maturational processes, this leads to potential measurement validity issues. The non-linear trajectories observed in adolescence could lend additional power to models utilizing longitudinal data, as larger changes occuring during specific age bands could help to increase model accuracy and performance.

In the current paper, we detail a novel BrainAGE model that utilizes longitudinal data, in the form of annualized change scores and random slopes of change. The data come from a large, longitudinal study of adolescent development, enabling both large sample sizes for training and metrics of structural change in the developing brain. By using metrics of change, we aim to take into account actual trajectories of development to more accurately assess brain maturation. In the following analyses, different modeling strategies are compared, including model type and change metric type. The accuracy and performance of the novel models are evaluated, and the best-performing model is used to relate predicted age gaps to pubertal development.

Methods

Participants

Participants were drawn from the Adolescent Brain Cognitive Development (ABCD) Study, a longitudinal, multisite study of over 10,000 adolescents from across the United States. Data came from the 5.1 release, which includes neuroimaging data from three waves, each collected two years apart. However, only partial data have been included for the third time point. At the time of recruitment, participants were 9-10 years old. Informed consent and assent was obtained from all participants and their families. For further details on recruitment procedures, see Garavan et al. (2018).

Only participants with high-quality scans were included, as determined by the imag_incl_t1w variable within the ABCD 5.1 Data Release. Scans were considered to be high quality if they

passed both manual quality control checks, and quality control checks during FreeSurfer preprocessing.

At baseline, 11,270 participants were recommended for inclusion (μ=9.92 ± 0.62 years old). At the two-year follow-up, 7,896 participants were recommended (μ=11.96 ± 0.65 years old), and at the four-year follow-up, 3,000 participants were recommended (μ=14.08 ± 0.69 years old). However, as this project took a longitudinal approach, only participants with data at multiple timepoints could be included.

Data Acquisition & Image Processing

Detailed imaging procedures can be found in Casey et al., 2019, and Hagler et al., 2019. In short, data were obtained from 21 sites using harmonized protocols. Structural scans were collected using one of three 3T scanner platforms (Phillips, Siemens, or GE) and a 32-channel head coil; 3-dimensional T1-weighted scans were collected with 1-mm voxel resolution.

As there is no standard for handling site and/or scanner differences in a BrainAGE framework, and past work has shown these models to be robust to site differences (Yu et al., 2023, preprint), we opted not to use site/scanner harmonization methods. However, we chose to exclude cortical thickness measurements from the model, as they are most susceptible to variability due to scanner differences (Fortin et al., 2018).

Change Score Calculation

Two approaches were used to calculate change scores, annualized change scores and a random slopes approach. These two approaches took the longitudinal nature of the data into account in different ways, enabling the evaluation of methodologies for calculating change in addition to modeling strategies.

Annualized Change

In the annualized change approach, raw change for each measure was first calculated separately for baseline - year 2 and year 2 - year 4. Raw change in each brain measure was then divided by the difference in age between the timepoints, which could vary by participant. This approach provided an annualized change measure for each region, for each participant. Data from both intervals were combined into the same dataframe, resulting in two columns for each brain region, one from each interval. To facilitate model comparisons, the participants used for the annualized change approach were limited to those available for the random slopes approach (N=2,446), described below.

Random Slopes

For the random slopes, data were first limited to participants with recommended imaging data available at all three timepoints, resulting in an N of 2,446. An unconditional growth model was fit for each brain measure, and random slopes for each participant were extracted. This resulted in participant-specific estimates of the slope of change for each of the 184 included features, described in further detail below. Model fitting was performed using the *lme()* function from R package *lme4*. Interview age and subject ID were included as random effects in the models.

Model Training

Model Features

Model features included cortical volume and surface area measurements, as well as bilateral global and subcortical volume measurements. Cortical volume and surface area measurements were obtained using the Desikan-Killiany atlas, while subcortical volume measurements were obtained using the default FreeSurfer ASEG atlas. In total, 184 regions were used for model training, corresponding to 184 features in the random slopes approach, and 368 features in the annualized change approach.

Training Procedures

Two types of models *(*elastic net and extreme gradient boosting/xgboost*)* were trained, with the goal of comparing modeling approaches. Both elastic net and xgboost approaches have been successfully used in prior adolescent BrainAGE research (Ball et al., 2021, Lewis et al., 2018, Drobinin et al., 2022). In all cases, age at the last time point (Year 4) was predicted by the features detailed above. All modeling was conducted using the tidymodels framework (Kuhn & Wickham, 2020).

Sample Split

Participants were split into training and testing sets, using an 80/20 split. 1,956 participants were included in the training sample, and 490 were included in the testing sample. During the splitting process, all members of the same family were randomly assigned to either the training or testing set. The same training and testing sets were used across both change approaches and model types.

Regularized Linear Regression

Regularized linear regression was conducted using glmnet() with an elastic net penalty, combining Ridge and Lasso penalties. Numeric variables were normalized/standardized using step_normalize(), while variables with zero or near-zero variance were removed using step_zv() and step_nzv(). Training was conducted using 10-fold cross-validation. During cross validation, members of the same family were again kept in either the training or test split.

Extreme Gradient Boosting

Extreme gradient boosting was conducted using the xgboost package (Chen & Guestrin, 2016). Variables with zero or near-zero variance were removed using step_zv() and step_nzv(). Training was conducted using 10-fold cross validation with 10 repeats. During cross validation, members of the same family were again kept in either the training or test split.

Analysis

Model Evaluation

The best-fitting model was determined as the model with the lowest mean absolute error (MAE). Across models, RMSE and R^2 are also reported.

Variable Importance
Variables with the highest contributions to each model were determined using the *vip* package (Greenwell & Boehmke, 2020).

Pubertal Development Metrics
Pubertal development was measured using the youth-report Pubertal Development Scale (PDS; Petersen et al., 1988). The PDS is a five-item measure of pubertal development, with three questions asked for all participants (height changes, skin changes, body hair growth), and two questions that varied based on the participant's assigned sex at birth. For female participants, these questions relate to the onset of menstruation and breast growth, and for male participants, they relate to facial hair and vocal changes. PDS scores were used only from the third time point, to align with the time point used for age prediction. Mean scores were used, with an average taken across all five items. Participants with incomplete data (N=19) were excluded from the analyses. Mean puberty scores were then correlated with BrainAGE.

Results
Model Performance
Performance metrics for the held-out test set for all model and change score calculation methods are described below, and summarized in *Table 1*. MAE and RMSE values are represented in terms of years.

Regularized Linear Regression - Elastic Net
The best-fitting elastic net model using annualized change scores used an alpha of 0.02 and a lambda of 0.316. The model performed with an MAE of 0.53, an RMSE of 0.63, and a R^2 of 0.18.

The best-fitting elastic net model using random slopes used an alpha of 0.14 and a lambda of 0.041. The model performed with a MAE of 0.52, a RMSE of 0.62, and a R^2 of 0.21.

Extreme Gradient Boosting
The best-fitting xgboost model using annualized change scores model performed with a MAE of 0.53, a RMSE of 0.63, and a R^2 of 0.15.

The best-fitting xgboost model using random slopes performed with a MAE of 0.53, a RMSE of 0.63, and a R^2 of 0.15.

Overall, the best-performing model was the elastic net model using the random slopes approach. Scatterplots of chronological versus predicted age for all models are shown in *Figure 1*.

Table 1. Performance Metrics

	Elastic Net		XGBoost	
	Annualized Change	Random Slopes	Annualized Change	Random Slopes
MAE	0.53	0.52	0.53	0.53
RMSE	0.63	0.62	0.63	0.63
R^2	0.18	0.21	0.15	0.15

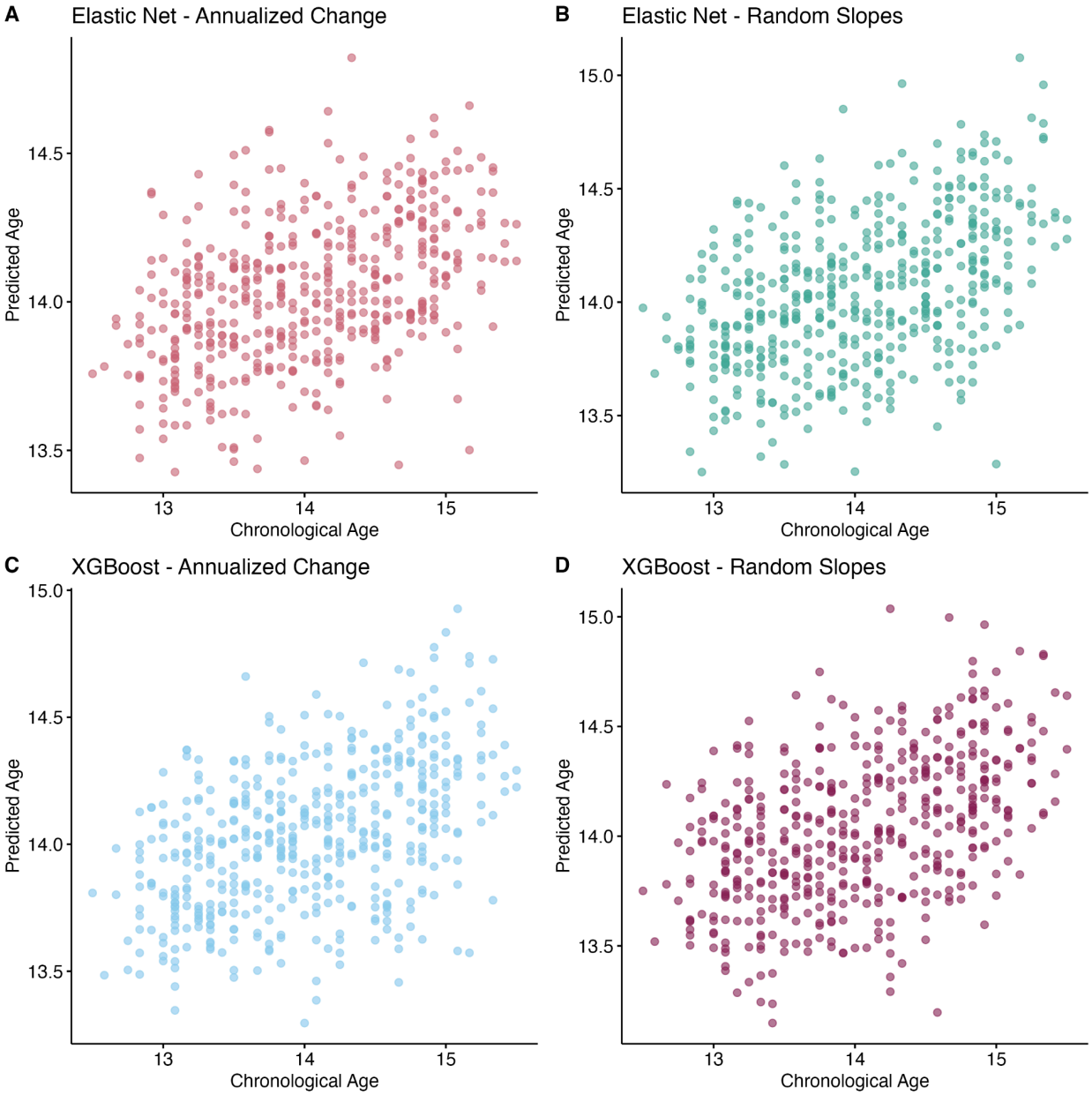


Figure 1. Chronological vs. predicted age for all models.

Feature Importance

Regularized Linear Regression

For the elastic net-random slopes approach, the highest-contributing features included the left precuneus area, right lingual volume, left amygdala volume, left supramarginal volume, and total left hemisphere surface area. *Figure 3* shows an atlas of the brain, with regions colored by their importance to the model predictions. For details on the highest-importance features to the other models, see the Appendix.

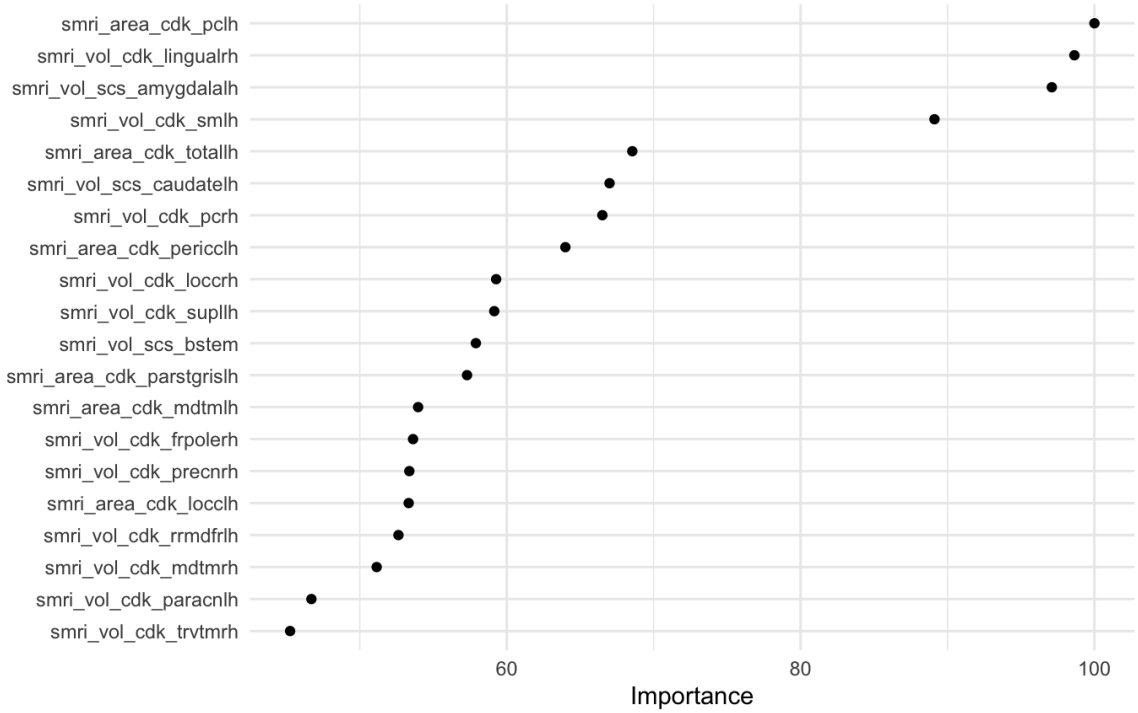


Figure 2. Ranking of the top 20 highest-contributing features.

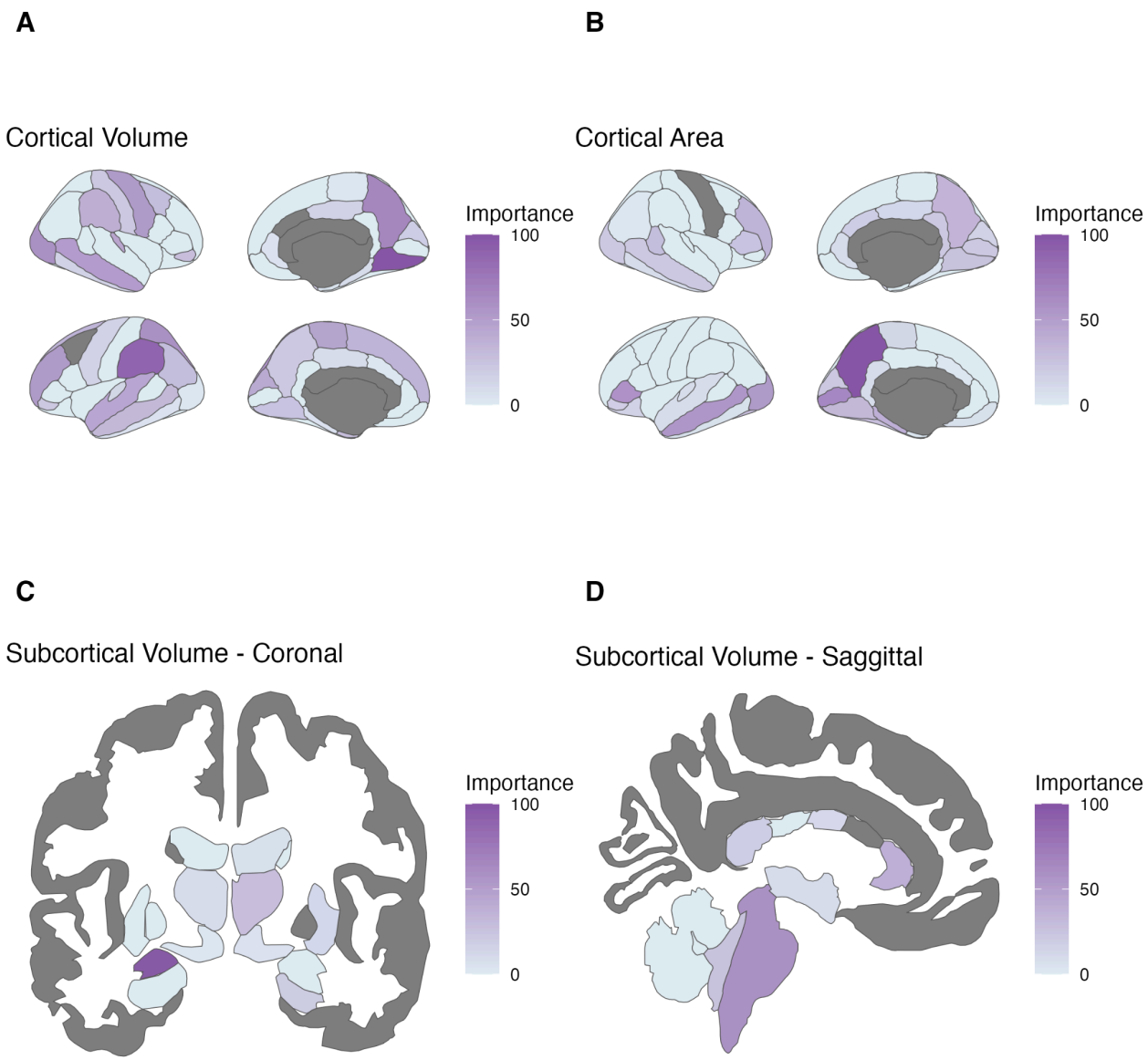


Figure 3. Variable importance and location within the brain. Regions with greater importance to the model predictions are shown in darker purple.

Pubertal Development Analysis

Pubertal development was negatively related to BrainAGE (p<.001, r = -0.19), such that participants with a predicted age older than their chronological age were less advanced in pubertal development. Distributions of BrainAGE by pubertal stage (categorized for visual purposes) are shown in *Figure 4*.

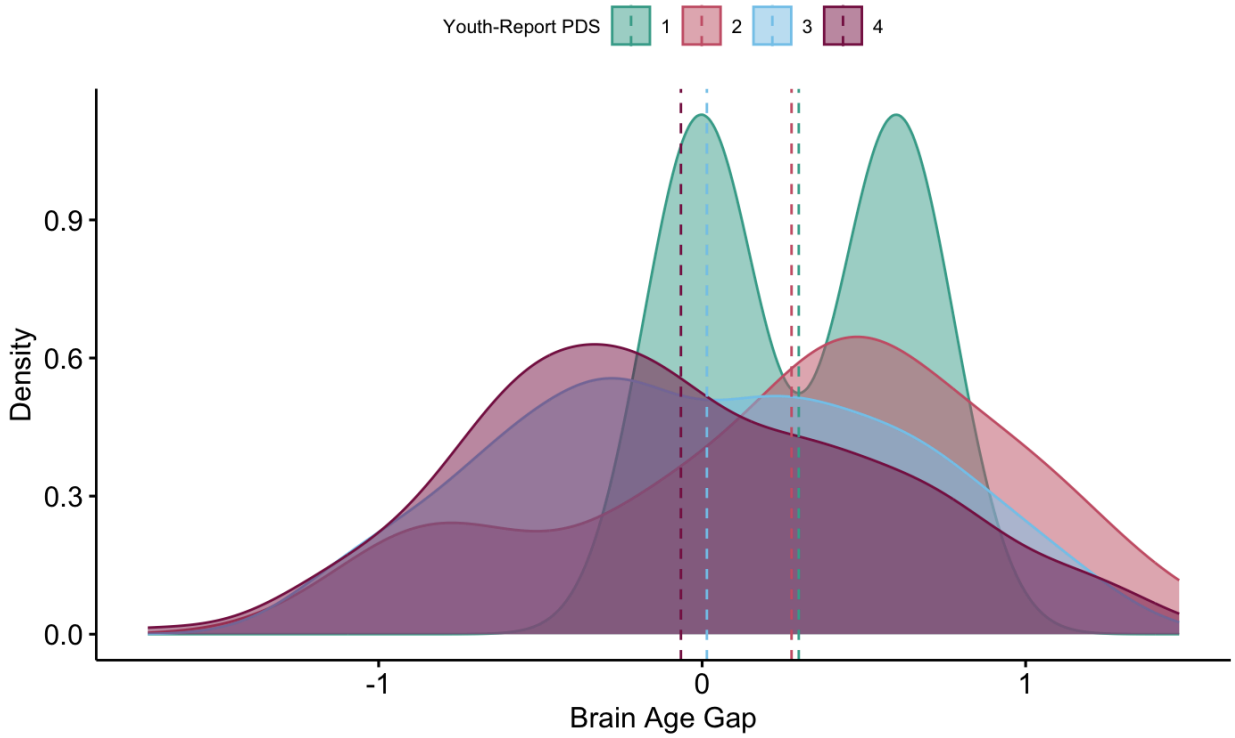


Figure 4. Distributions of BrainAGE by pubertal stage. Vertical dashed lines represent the mean BrainAGE for each pubertal category.

Discussion

Model Performance

Overall, the best-fitting model was the elastic net model using random slopes. However, all four models performed almost identically. MAEs and RMSEs were essentially the same across the remaining three models, and the best performing model was an improvement of .01 for both metrics. R^2 varied more across the models, but remained low, ranging from 0.15 to 0.21. The two models trained using the xgboost framework resulted in the lowest R^2, both at 0.15. Importantly, while the MAEs for the models were quite low compared to existing BrainAGE models, the models explained only a small amount of variance, as measured by R^2. It's possible that utilizing a larger test sample, or testing on an external dataset could help to further characterize performance differences between the models.

Prior work using longitudinal measures of brain change to predict BrainAGE is limited. The only published work to do so comes from Wierenga et al. (2019), who used random forests to create a model based on both cross-sectional and longitudinal measures, resulting in a MAE of 1.976. In their study, the model with both cross-sectional and longitudinal measures was an improvement over a solely cross-sectional model, which had an MAE of 2.01. Their combined model had an adjusted R^2 of .79, significantly better than the best-performing model reported here. There are a few potential sources for this discrepancy, including the age range covered in training, the method of calculating change, model features, and the inclusion of both

cross-sectional and longitudinal measures in the Wierenga model. Interestingly, the Wierenga model used a much smaller sample size (n=198) for training, so the reduced R^2 in the current study was an unexpected result.

Existing cross-sectional models are more common, and often perform with MAEs in the 1-2 year range (Franke & Gaser, 2019). Metrics on model performance in terms of R^2 are less commonly reported, but some studies have reported R^2 values of 0.22, 0.49, 0.79, and 0.81, showing a wide variation (Dehestani et al., 2023, Ball et al., 2021, Keding et al., 2021, Kelly et al., 2022). Additionally, the fact that many studies solely report MAE as a fit metric is a source for concern, as it's possible that models may have a low MAE, but also a low r/R^2, as was the case in the current study. Recent work has also shown that BrainAGE models using a narrower age range tend to have lower r and R^2 values (de Lange et al., 2022). This may have been a factor in the current study, and aligns with the similarly low R^2 value of 0.22 found in Dehestani et al. (2023), which utilized a similarly small age range from the same sample, though covering a younger range.

Feature Contributions

The highest-contributing features were primarily cortical volume measurements, though the overall highest-contributing feature was the surface area of the left precuneus, and there were a few subcortical volume measures as high contributing features.

In terms of interpretations of these features, structural changes in many of the identified regions have been related to pubertal development in adolescence, which is known to be a major driver of adolescent brain development (Vijayakumar et al., 2018). For the specific regions identified as high contributors, past work has characterized some of the changes occurring in these regions, which may be helpful for understanding why they have emerged as important features.

For the precuneus, little work exists that characterizes surface area changes, but past work has found volume reductions during adolescence, which have been tied to pubertal maturation (Sullivan et al., 2011). Additionally, lingual volume is known to decrease during the age range examined in the current study (Tamnes et al., 2017). Amygdala volume also changes during this time, and has been found to increase in volume, in contrast to many other regions exhibiting volume reductions (Østby et al., 2009). Similar to the precuneus, amygdala development has been related to pubertal maturation (Hu et al., 2013). For the supramarginal gyrus, nonlinear changes have been identified, and changes in density have been related to estradiol, a key hormone during puberty (Sussman et al., 2016, Peper et al., 2009).

Pubertal Development

The negative relationship between BrainAGE and pubertal development found here contrasts the existing literature, which has found cross-sectional BrainAGE to be positively related to pubertal development (Dehestani et al., 2022, Holm et al., 2023). However, there are a few notable differences between these studies. The existing work relied on cross-sectional BrainAGE for training, though Holm and colleagues used longitudinal data for testing to investigate changes in BrainAGE over time. Both studies utilized a slightly younger age range

than that of the current study, as they used data from the ABCD Study, but were limited to the first two timepoints by the data release schedule.

At face value, it seems counterintuitive that participants who were more advanced in puberty had a lower BrainAGE than those less advanced, given that pubertal development is known to have strong influences on brain development (Vijayakumar et al., 2018). However, in looking at the distributions of pubertal stage and BrainAGE, it appears that the relationship is a bit more complicated. Those in the later stages of puberty had an average BrainAGE closer to 0, indicating their predicted and chronological ages were similar. Meanwhile, those in the earlier stages of puberty were predicted to be older. Given that the participants in this study were between the ages of 13-15, it may be that those who had not yet hit puberty or had only just begun pubertal development by mid-adolescence are experiencing a different developmental pattern than those following more normative pubertal timing. Additional work is needed to further investigate this finding, and untangle the potentially unique influences of pubertal stage and timing of pubertal onset.

Limitations

While our ability to compare approaches across methods was improved by limiting participants to only those with three timepoints of data available, this approach did limit the available training sample. While the training sample of 1,956 participants is larger than many samples used in the adolescent BrainAGE literature, it would be ideal to make full use of the sample. Additionally, there may be systematic differences in the participants that had usable data available at all three waves compared to those that did not. Past studies have identified statistically significant differences in household income, parental education, race/ethnicity, and age in the participants recommended for inclusion at the first wave of ABCD data collection (Chaarani et al., 2021). Additionally, excessive subject motion is a common cause of low scan quality, and in-scanner motion is related to factors such as age, where older adolescents tend to have less motion.

Future Directions

It's possible that other methods of calculating change could affect or improve model performance, including annualized percent change. Annualized percent change could further account for individual differences in overall brain size by examining change as a percent of the total measure.

Additionally, annualized change or percent change could be used in such a way that age is predicted from just the previous time interval (i.e. predicting Time2 from Time1->Time2 and Time3 from Time2->Time3), rather than predicting age at the latest time point from all prior intervals. While providing data about a smaller window of change, this method could potentially provide different information by widening the range of predicted ages. It will also likely be beneficial to integrate additional windows of change as more timepoints of data become available from the ABCD Study, as this may help to more accurately characterize unique changes at different windows in development. Given that narrower age ranges typically result in smaller R^2 values, as discussed in *Model Performance*, widening the outcome age range with additional waves of data or alternate approaches could be beneficial.

Additionally, future work should further explore the relationship between change score BrainAGE and changes in maturational metrics, such as pubertal and cognitive development, and outcomes that have been hypothesized to be related to altered neurodevelopmental trajectories. Doing so would be an important step in evaluating both the conceptual validity and practical applications of this methodology.

Conclusion

In summary, we were able to create a novel BrainAGE model that utilized longitudinal data for model training. By comparing multiple models and training methods, an elastic net model trained using individualized slopes of change for each brain region emerged as the best-performing model. However, the results of this study should be taken with a note of caution, as all models tested performed quite similarly, and the variance explained by the models was relatively low. BrainAGE as predicted by the best-performing model was found to be negatively correlated with pubertal development, which necessitates additional exploration. In the future, more work should be done to validate the model with regards to behavioral outcomes and relationships to maturational metrics in adolescence. However, the creation of this model provides an important starting point for future work to build on.

References

Ball, G., Kelly, C. E., Beare, R., & Seal, M. L. (2021). Individual variation underlying brain age estimates in typical development. *NeuroImage*, *235*, 118036. https://doi.org/10.1016/j.neuroimage.2021.118036

Brown, T. T., Kuperman, J. M., Chung, Y., Erhart, M., McCabe, C., Hagler, D. J., 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., … Dale, A. M. (2012). Neuroanatomical Assessment of Biological Maturity. *Current Biology*, *22*(18), 1693–1698. https://doi.org/10.1016/j.cub.2012.07.002

Brown, T. T., Kuperman, J. M., Chung, Y., Erhart, M., McCabe, C., Hagler, D. J., 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., … Dale, A. M. (2012). Neuroanatomical Assessment of Biological Maturity. *Current Biology*, *22*(18), 1693–1698. https://doi.org/10.1016/j.cub.2012.07.002

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C. A., Wager, T. D., Banich, M. T., Speer, N. K., Sutherland, M. T., Riedel, M. C., Dick, A. S., Bjork, J. M., Thomas, K. M., … Dale, A. M. (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, *32*, 43–54. https://doi.org/10.1016/j.dcn.2018.03.001

Chaarani, B., Hahn, S., Allgaier, N. *et al.* Baseline brain function in the preadolescents of the ABCD Study. *Nat Neurosci* 24, 1176–1186 (2021). https://doi.org/10.1038/s41593-021-00867-9

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939785

Cohen, J. W., Ramphal, B., DeSerisy, M., Zhao, Y., Pagliaccio, D., Colcombe, S., Milham, M. P., & Margolis, A. E. (2023). Relative brain age is associated with socioeconomic status and anxiety/depression problems in youth. *Developmental Psychology*, https://doi.org/10.1037/dev0001593

Cropley, V. L., Tian, Y., Fernando, K., Mansour L., S., Pantelis, C., Cocchi, L., & Zalesky, A. (2021). Brain-Predicted Age Associates With Psychopathology Dimensions in Youths. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *6*(4), 410–419. https://doi.org/10.1016/j.bpsc.2020.07.014

Dehestani, N., Whittle, S., Vijayakumar, N., & Silk, T. J. (2023). Developmental brain changes during puberty and associations with mental health problems. *Developmental Cognitive Neuroscience*, *60*, 101227. https://doi.org/10.1016/j.dcn.2023.101227

de Lange, A.-M. G., Anatürk, M., Rokicki, J., Han, L. K. M., Franke, K., Alnæs, D., Ebmeier, K. P., Draganski, B., Kaufmann, T., Westlye, L. T., Hahn, T., & Cole, J. H. (2022). Mind the gap: Performance metric evaluation in brain-age prediction. *Human Brain Mapping*, *43*(10), 3113–3129. https://doi.org/10.1002/hbm.25837

Drobinin, V., Van Gestel, H., Helmick, C. A., Schmidt, M. H., Bowen, C. V., & Uher, R. (2022). The Developmental Brain Age Is Associated With Adversity, Depression, and Functional Outcomes Among Adolescents. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *7*(4), 406–414. https://doi.org/10.1016/j.bpsc.2021.09.004

Fortin, J. P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., ... & Shinohara, R. T. (2018). Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*, *167*, 104-120.

Franke, K., & Gaser, C. (2012). Longitudinal Changes in Individual BrainAGE in Healthy Aging, Mild Cognitive Impairment, and Alzheimer's Disease. *GeroPsych*, *25*(4), 235–245. https://doi.org/10.1024/1662-9647/a000074

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

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., Jernigan, T., Potter, A., Thompson, W., & Zahs, D. (2018). Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience*, *32*, 16–22. https://doi.org/10.1016/j.dcn.2018.04.004

Greenwell, Brandon M., and Bradley C. Boehmke. 2020. “Variable Importance Plots—An Introduction to the vip Package.” *The R Journal* 12 (1): 343–66. https://doi.org/10.32614/RJ-2020-013.

Hagler, D. J., Hatton, SeanN., Cornejo, M. D., Makowski, C., Fair, D. A., Dick, A. S., Sutherland, M. T., Casey, B. J., Barch, D. M., Harms, M. P., Watts, R., Bjork, J. M., Garavan, H. P., Hilmer, L., Pung, C. J., Sicat, C. S., Kuperman, J., Bartsch, H., Xue, F., … Dale, A. M. (2019). Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *NeuroImage*, *202*, 116091. https://doi.org/10.1016/j.neuroimage.2019.116091

Holm, M. C., Leonardsen, E. H., Beck, D., Dahl, A., Kjelkenes, R., de Lange, A.-M. G., & Westlye, L. T. (2023). Linking brain maturation and puberty during early adolescence using longitudinal brain age prediction in the ABCD cohort. *Developmental Cognitive Neuroscience*, *60*, 101220. https://doi.org/10.1016/j.dcn.2023.101220

Hu S, Pruessner JC, Coupé P, Collins DL. (2013). Volumetric analysis of medial temporal lobe structures in brain development from childhood to adolescence. *NeuroImage.* 2013;74C:276–287.

Keding, T. J., Heyn, S. A., Russell, J. D., Zhu, X., Cisler, J., McLaughlin, K. A., & Herringa, R. J. (2021). Differential Patterns of Delayed Emotion Circuit Maturation in Abused Girls With and Without Internalizing Psychopathology. *American Journal of Psychiatry*, *178*(11), 1026–1036. https://doi.org/10.1176/appi.ajp.2021.20081192

Kelly, C., Ball, G., Matthews, L. G., Cheong, J. L., Doyle, L. W., Inder, T. E., Thompson, D. K., & Anderson, P. J. (2022). Investigating brain structural maturation in children and adolescents born very preterm using the brain age framework. *NeuroImage*, *247*, 118828. https://doi.org/10.1016/j.neuroimage.2021.118828

Kuhn M, Wickham H (2020). *Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles.* https://www.tidymodels.org.

Mills, K. L., Goddings, A. L., Herting, M. M., Meuwese, R., Blakemore, S. J., Crone, E. A., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., & Tamnes, C. K. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage*, *141*, 273–281. https://doi.org/10.1016/j.neuroimage.2016.07.044

Mills, K. L., Siegmund, K. D., Tamnes, C. K., Ferschmann, L., Wierenga, L. M., Bos, M. G. N., Luna, B., Li, C., & Herting, M. M. (2021). Inter-individual variability in structural brain development from late childhood to young adulthood. *NeuroImage*, *242*, 118450. https://doi.org/10.1016/j.neuroimage.2021.118450

Ostby, Y., Tamnes, C. K., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., & Walhovd, K. B. (2009). Heterogeneity in subcortical brain development: A structural magnetic resonance imaging study of brain maturation from 8 to 30 years. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *29*(38), 11772–11782. https://doi.org/10.1523/JNEUROSCI.1242-09.2009

Peper, J. S., Brouwer, R. M., Schnack, H. G., van Baal, G. C., van Leeuwen, M., van den Berg, S. M., Delemarre-Van de Waal, H. A., Boomsma, D. I., Kahn, R. S., & Hulshoff Pol, H. E. (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

Petersen, A. C., Crockett, L., et al. (1988) A self-report measure of pubertal status: Reliability, validity, and initial norms. J Youth Adolesc 17(2): 117-133.

Shaw, P., Gogtay, N., & Rapoport, J. (2010). Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. *Human Brain Mapping*, *31*(6), 917–925. https://doi.org/10.1002/hbm.21028

Sullivan, E. V., Pfefferbaum, A., Rohlfing, T., Baker, F. C., Padilla, M. L., & Colrain, I. M. (2011). Developmental change in regional brain structure over 7 months in early adolescence: comparison of approaches for longitudinal atlas-based parcellation. *NeuroImage*, *57*(1), 214–224. https://doi.org/10.1016/j.neuroimage.2011.04.003

Sussman, D., Leung, R. C., Chakravarty, M. M., Lerch, J. P., & Taylor, M. J. (2016). The developing human brain: age-related changes in cortical, subcortical, and cerebellar anatomy. *Brain and behavior*, *6*(4), e00457. https://doi.org/10.1002/brb3.457

Tamnes, C. K., Herting, M. M., Goddings, A. L., Meuwese, R., Blakemore, S. J., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., Crone, E. A., & Mills, K. L. (2017). Development of the Cerebral Cortex across Adolescence: A Multisample Study of Inter-Related Longitudinal Changes in Cortical Volume, Surface Area, and Thickness. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *37*(12), 3402–3412. https://doi.org/10.1523/JNEUROSCI.3302-16.2017

Vijayakumar, N., Op de Macks, Z., Shirtcliff, E. A., & Pfeifer, J. H. (2018). Puberty and the human brain: Insights into adolescent development. *Neuroscience & Biobehavioral Reviews*, *92*, 417–436. https://doi.org/10.1016/j.neubiorev.2018.06.004

Wierenga, L. M., Bos, M. G. N., van Rossenberg, F., & Crone, E. A. (2019). Sex Effects on Development of Brain Structure and Executive Functions: Greater Variance than Mean Effects. *Journal of Cognitive Neuroscience*, *31*(5), 730–753. https://doi.org/10.1162/jocn_a_01375

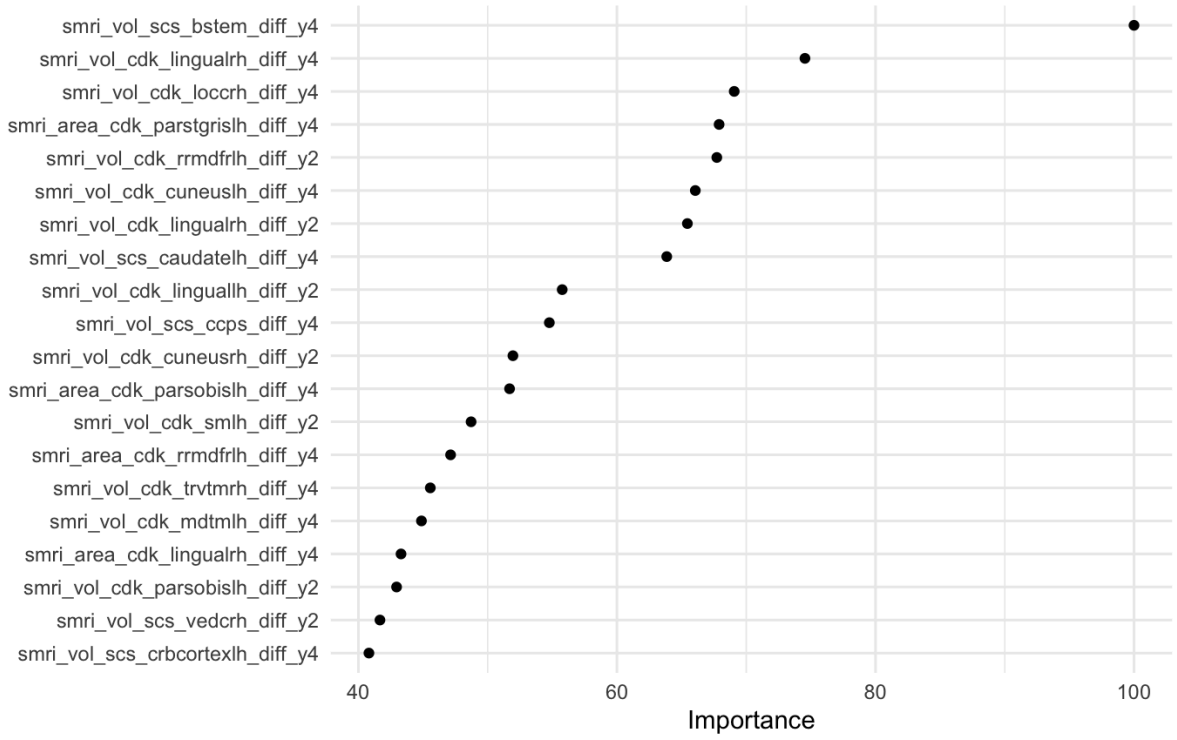
Yu, Y., Cui, H.-Q., Haas, S. S., New, F., Sanford, N., Yu, K., Zhan, D., Yang, G., Gao, J.-H., Wei, D., Qiu, J., Bernhardt, B., Thompson, P., Frangou, S., Ge, R., & Center, E. W. A. (2023). *Brain-Age Prediction: Systematic Evaluation of Site Effects, and Sample Age Range and Size* (p. 2023.11.06.565917). bioRxiv. https://doi.org/10.1101/2023.11.06.565917

Appendix

Variable Importance

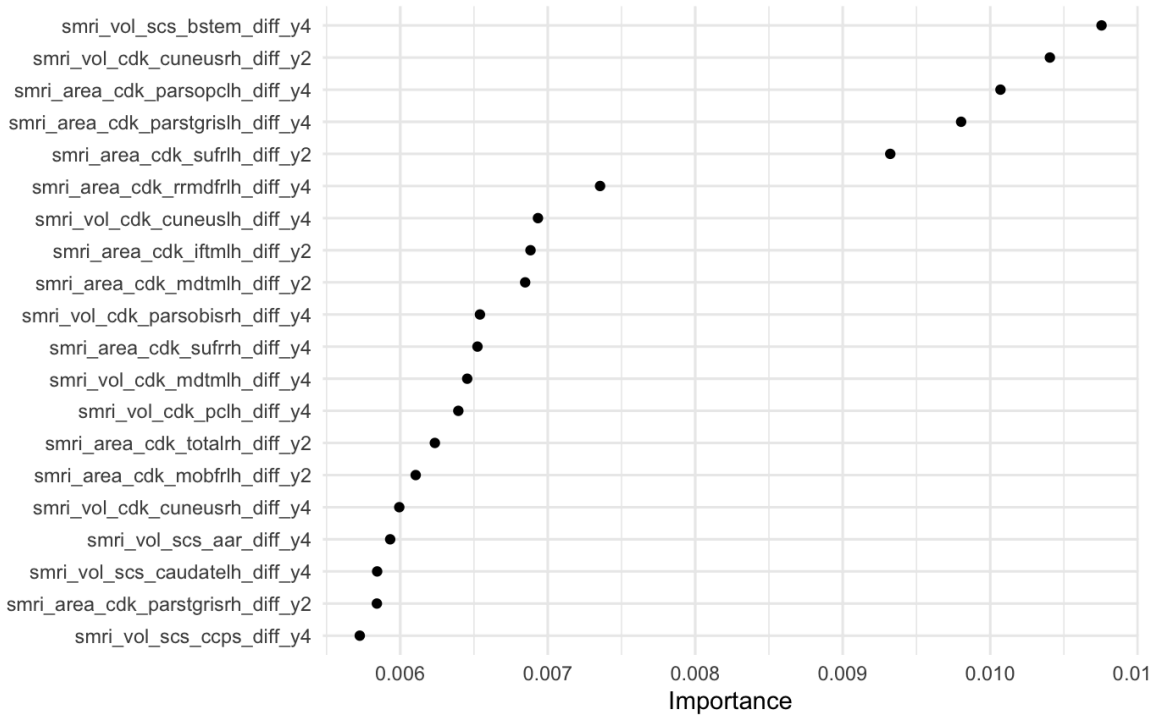
Regularized Linear Regression

For the annualized change score approach, the highest-contributing features included the brainstem, right lingual, right locus coeruleus, and left pars triangularis volumes from the second time band (year 2 - year 4).

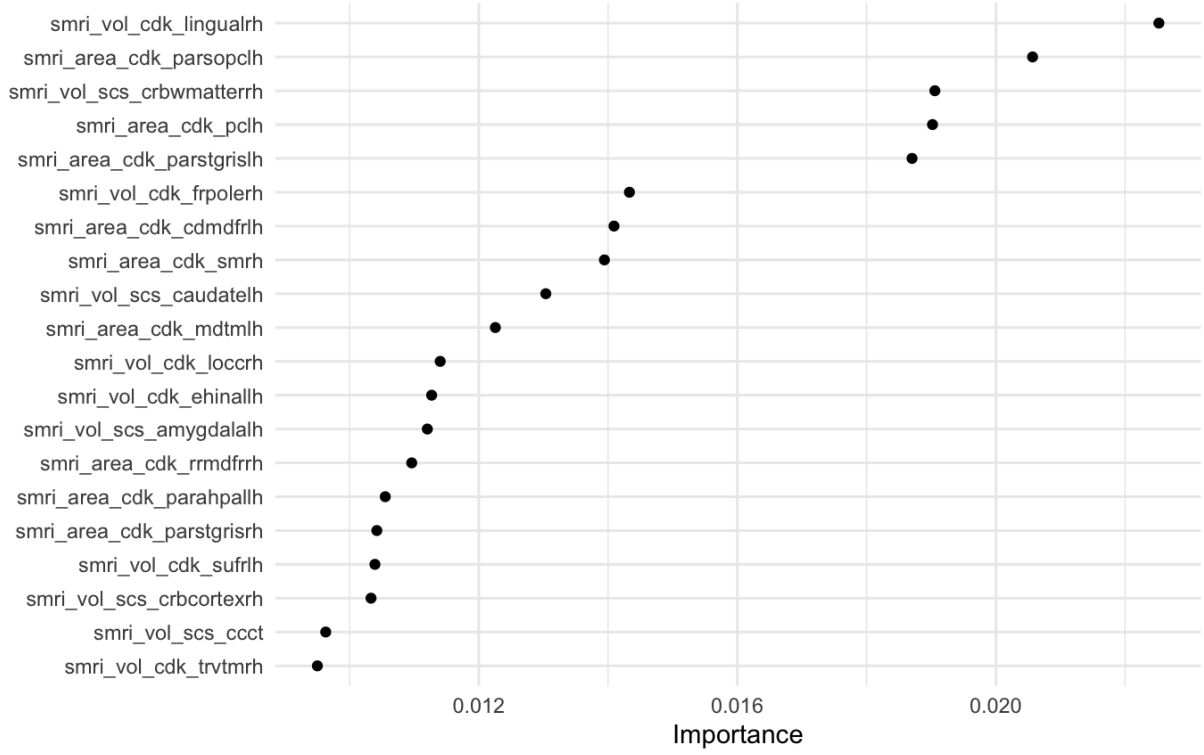


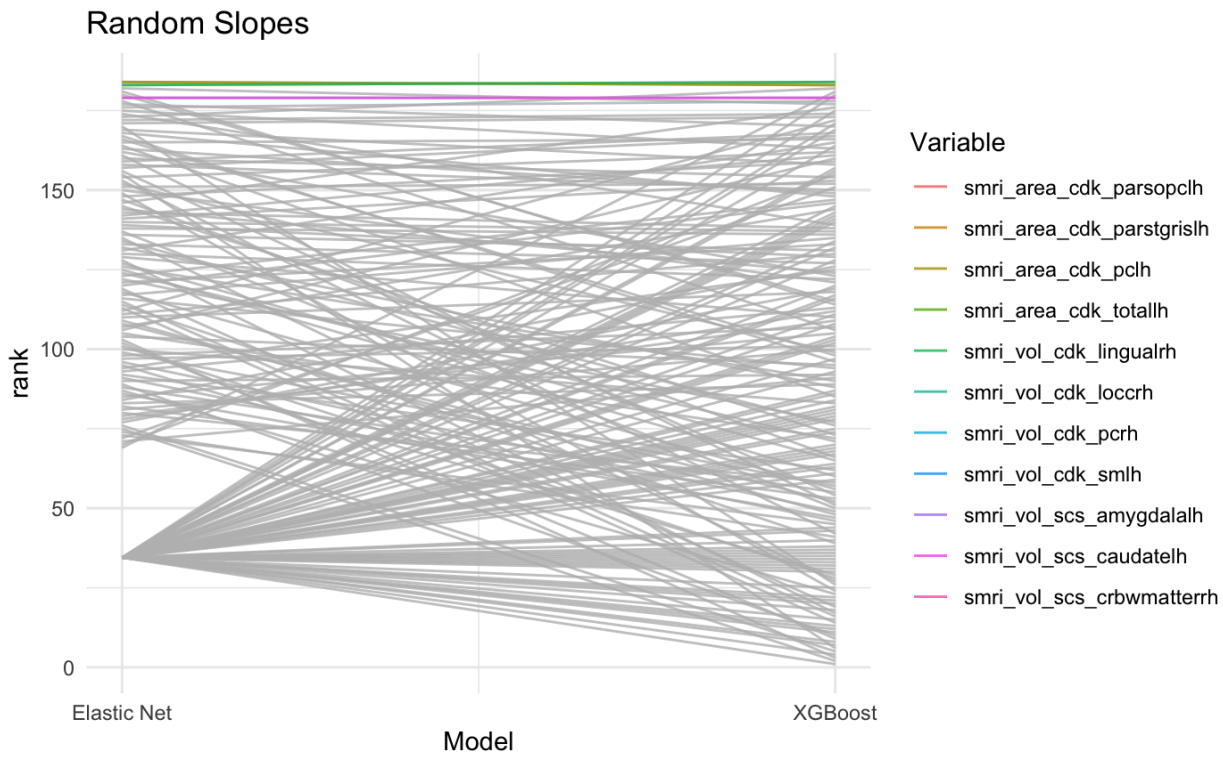
Extreme Gradient Boosting

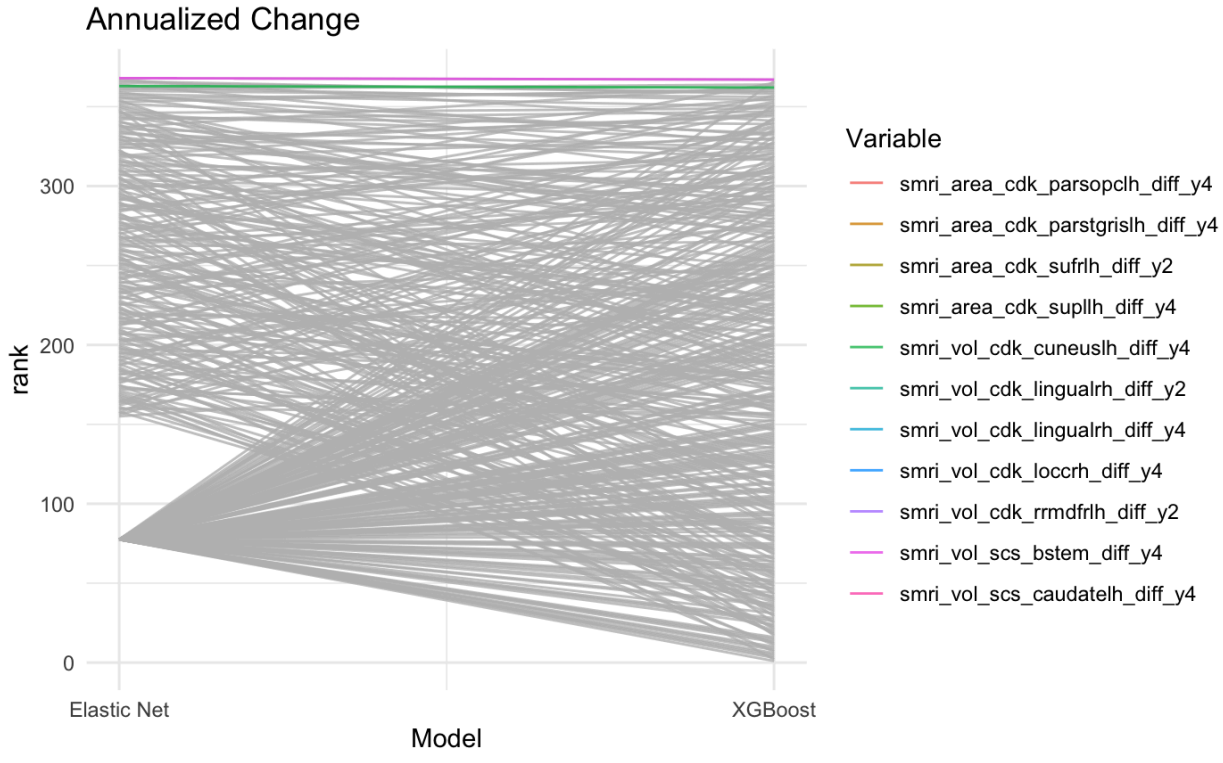
For the annualized change score approach, the highest-contributing features were brainstem volume, right cuneus volume, left pars opercularis surface areas, left pars triangualris surface area, and left superior frontal surface area.



For the random slopes approach, the highest-contributing features were the right lingual volume, left pars opercularis volume, right cerebral white matter volume, and left precuneus area.




Random Slopes
rank
150
100
50
0
Elastic Net
XGBoost
Model
Variable
smri_area_cdk_parsopclh
smri_area_cdk_parstgrislh
smri_area_cdk_pclh
smri_area_cdk_totallh
smri_vol_cdk_lingualrh
smri_vol_cdk_loccrh
smri_vol_cdk_pcrh
smri_vol_cdk_smlh
smri_vol_scs_amygdalalh
smri_vol_scs_caudatelh
smri_vol_scs_crbwmatterrh


Annualized Change
rank
300
200
100
0
Elastic Net
XGBoost
Model
Variable
smri_area_cdk_parsopclh_diff_y4
smri_area_cdk_parstgrislh_diff_y4
smri_area_cdk_sufrlh_diff_y2
smri_area_cdk_suplh_diff_y4
smri_vol_cdk_cuneuslh_diff_y4
smri_vol_cdk_lingualrh_diff_y2
smri_vol_cdk_lingualrh_diff_y4
smri_vol_cdk_loccrh_diff_y4
smri_vol_cdk_rrmdfrlh_diff_y2
smri_vol_scs_bstem_diff_y4
smri_vol_scs_caudatelh_diff_y4