

RESEARCH

Open Access



Menarche, pubertal timing and the brain: female-specific patterns of brain maturation beyond age-related development

Nina Gottschewsky^{1,4*†}, Dominik Kraft^{1†}  and Tobias Kaufmann^{1,2,3*} 

Abstract

Background Puberty depicts a period of profound and multifactorial changes ranging from social to biological factors. While brain development in youths has been studied mostly from an age perspective, recent evidence suggests that pubertal measures may be more sensitive to study adolescent neurodevelopment, however, studies on pubertal timing in relation to brain development are still scarce.

Methods We investigated if pre- vs. post-menarche status can be classified using machine learning on cortical and subcortical structural magnetic resonance imaging (MRI) data from strictly age-matched adolescent females from the Adolescent Brain Cognitive Development (ABCD) cohort. For comparison of the identified menarche-related patterns to age-related patterns of neurodevelopment, we trained a brain age prediction model on data from the Philadelphia Neurodevelopmental Cohort and applied it to the same ABCD data, yielding differences between predicted and chronological age referred to as brain age gaps. We tested the sensitivity of both these frameworks to measures of pubertal maturation, specifically age at menarche and puberty status.

Results The machine learning model achieved moderate but statistically significant accuracy in the menarche classification task, yielding for each subject a class probability ranging from 0 (pre-) to 1 (post- menarche). Comparison to brain age predictions revealed shared and distinct patterns of neurodevelopment captured by both approaches. Continuous menarche class probabilities were positively associated with brain age gaps, but only the menarche class probabilities—not the brain age gaps—were associated with age at menarche.

Conclusions This study demonstrates the use of a machine learning model to classify menarche status from structural MRI data while accounting for age-related neurodevelopment. Given its sensitivity towards measures of puberty timing, our work suggests that menarche class probabilities may be developed toward an objective brain-based marker of pubertal development.

Highlights

- We classified pre- vs. post-menarche status in adolescent females from structural brain imaging data

[†]Nina Gottschewsky and Dominik Kraft contributed equally to this work.

*Correspondence:

Nina Gottschewsky
nina.gottschewsky@maxplanckschools.de
Tobias Kaufmann
tobias.kaufmann@med.uni-tuebingen.de

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

- We compared class probabilities to brain-age predictions to disentangle puberty- vs. age-related patterns of brain development
- The derived continuous brain-based menarche class probabilities captured shared but also unique variations of adolescent neurodevelopment, and were associated with pubertal timing and status

Keywords Female brain development, Menarche, Pubertal timing, Machine learning on imaging data

Plain language summary

Puberty is a period of substantial changes in the life of youths, and these include profound brain changes. Most studies have investigated age related changes in brain development, recent work however suggests that looking at brain development through the lens of pubertal development can provide additional insights beyond age effects. We here analyzed brain imaging data from a group of same-aged adolescent girls from the Adolescent Brain Cognitive Development study. Our goal was to investigate if we could determine from brain images whether a girl had started her menstrual period (menarche) or not, and we used machine learning to classify between them. This machine learning model does not just return a “yes/no” decision, but also returns a number between 0 and 1 indicating a probability to be pre- (0) or post- (1) menarche. To rule out that our approach only maps age-related development, we selected a strictly age-matched sample of girls and compared our classification model to a brain age model trained on independent individuals. Our model classified between pre- and post-menarche with moderate accuracy. The obtained class probability was partly related to age-related brain development, but only the probability was significantly associated with pubertal timing (age at menarche). In summary, our study uses a machine learning model to estimate whether a girl has reached menarche based on her brain structure. This approach offers new insights into the connection between puberty and brain development and might serve as an objective way to assess pubertal timing from imaging data.

Background

Adolescence is a time of profound changes to the body and the brain, with substantial impact on an individual's behaviour, emotions, and self-perception, among other things [1]. This transition includes puberty, the time period during which an individual acquires the capability for sexual reproduction [2]. The latter is characterised by an interplay of gonadotropin-releasing hormone, gonadotropins such as follicle-stimulating hormone and luteinizing hormone, and sex hormones such as androgens, estrogens and progesterones. Together, they not only drive changes of the body, but also directly act on the brain [3]. Studies using magnetic resonance imaging (MRI) to investigate human brain anatomy have illustrated that the brain undergoes significant changes during adolescence described by a complex yet orchestrated interplay of progressive (e.g., myelination) and regressive (e.g., pruning) neuronal processes [4]. While brain development in youths has been commonly investigated through the lens of age-related brain maturation, there has been an increasing number of studies focusing on neurodevelopment mediated by pubertal processes in youth [5–7]. These studies suggest that puberty-related brain development cannot be simply explained by age trajectories but rather goes beyond the effects of growing older [7–9] and consequently that pubertal development may thus

be a more sensitive measure to study neurodevelopment in youths as compared to age [6].

A recent systematic review on the relationship between pubertal and structural brain development in human adolescents describes brain wide reductions in cortical grey matter thickness and volume associated with progressed pubertal maturation from both cross-sectional and longitudinal studies [7]. Findings suggest that these effects are global across the brain with frontal regions showing the most pronounced effects [7, 10]. Alongside cortical changes, advanced pubertal maturation is also associated with subcortical brain development, in particular the amygdala and hippocampus [11]. Across studies these effects are subject to sex differences, which not only manifest in varying effect sizes but also sometimes in opposing effect directions in males and females [5, 12].

While methodological choices, such as accounting for age in statistical models, may factor into the diverging observations, these differences may also stem from variability that is inherent to pubertal maturation [13]. Although all individuals pass through the same pubertal stages, there is large variability regarding pubertal timing and the speed of progression [14, 15]. Pubertal timing describes the time point at which an individual reaches certain pubertal milestones in comparison to their peers of the same age [16]. While pubertal timing appears to be highly heritable [17, 18], recent evidence is also linking

variation in pubertal timing to environmental factors, such as nutrition intake, socioeconomic status or obesity [14]. This malleability may consequently lead to pubertal onsets that deviate in their timing and individuals thus experiencing early or late pubertal onsets [5]. Interestingly these deviations in pubertal timing appear to be associated with physical and psychiatric health issues [5, 19].

Many studies over the years have shown an association between pubertal timing and psychopathology [20–22]. In boys, evidence concerning the effect of pubertal timing on health risks is inconsistent and could be best described by the ‘off-time hypothesis’, that is either very early or very late onset [23]. In contrast, evidence for the association between health risk and pubertal timing in girls has been well-replicated, converging on the so-called ‘early timing hypothesis’, which posits that early maturing girls (most often assessed using age at menarche as a proxy measure, i.e. age at which individuals experience their first menstruation) are more likely to experience adverse mental health outcomes than their on time and late maturing peers [24, 25]. Therefore, pubertal timing and its malleability depict a critical tipping point which may set the course for later vulnerability and worse (mental) health outcome.

While most puberty-related imaging studies to date have focused on investigating the association between the brain and puberty status (i.e. the quantification of pubertal characteristics indicating a more or less advanced maturation akin to the transition through pubertal stages), imaging studies on pubertal timing—despite its importance for emerging (mental) health risks—are to the best of our knowledge scarce. The current study investigated the impact of pubertal timing on brain maturation, deploying age-matching to control for age-related neurodevelopment. Using structural imaging data from the Adolescent Brain Cognitive Development cohort (ABCD; [26]) we aimed at classifying pre- and post-menarcheal females using a machine learning model. To validate the sensitivity of our approach and to test the biological validity of the obtained class probabilities, we drew comparison to a brain age prediction framework, investigating to what extent both approaches capture the same or distinct neurodevelopmental variance in the female adolescent brain.

Methods

Sample descriptions

ABCD: For the menarche classification and as the test sample for the age prediction model, we included data of $N=3248$ female (henceforth referring to individuals assigned female at birth; mean age=11.91 years, $SD=0.65$) participants of the Adolescent Brain Cognitive

Development study 2-year follow up data [26]. Study protocols have been approved by either local institutional review boards (IRB) or by reliance agreements with the central IRB at University of California. For each study participant, structural brain imaging features were obtained from the tabulated imaging data provided by the ABCD release 4.0 [27]. The 2-year follow up data was chosen because it offers the most balanced distribution of pre- and post-menarcheal individuals. Subjects with missing MRI or missing relevant demographic data were excluded. Furthermore, those who did not answer either ‘yes’ or ‘no’ to the question ‘Have you begun to menstruate (started to have your period)?’ from the ABCD Youth Pubertal Development Scale and Menstrual Cycle Survey History (PDMS) [28], or whose imaging data quality was deemed too low for inclusion by two ABCD researchers, were excluded (see Additional file 1: Methods for a detailed description of the in- and exclusion procedure). From the PDMS data we determined pubertal status ranging from prepubertal to postpubertal. In brief, we summed pubic hair growth and breast development scores and incorporated information about menarche, and converted the resulting score to a pubertal status category according to a scheme provided by the ABCD study (variable: `pds_p_ss_female_category`). Pubertal status was calculated from youth-reported as well as caregiver-reported data to account for differences in the perception of pubertal maturation.

PNC: We used data from $N=786$ female participants (mean age=15.25 years, $SD=3.65$) of the Philadelphia Neurodevelopmental Cohort (PNC; [29]) as an independent training sample to derive an age prediction model. In PNC, all study procedures were approved by the respective institutional review boards. We processed the T1 MRI images using FreeSurfer (version 7.1.1) [30] and derived the same cortical and subcortical features as used for the ABCD cohort. Euler numbers were used as a proxy of image quality for quality control [31]. Subjects with missing MRI, missing demographic data, a Euler number more than three standard deviations below the mean, or those with a medical rating of 3 or higher (severe medical condition) were excluded.

MRI data description

For each subject in both data sets we included a total number of 234 anatomical MRI features. Specifically, we used 30 subcortical features as well as, for each hemisphere, 34 volume, 34 thickness, and 34 area cortical features matching the Desikan–Killiany atlas [32] (see Additional file 1: Table S1). Of note, since ABCD data was acquired across 21 study sites, we performed batch harmonization with *neuroCombat* (v.0.2.12) [33]

for each individual modality and training and test sets independently.

Statistical analyses

All statistical analyses were performed in python (v.3.11.5) [34]. Basic data handling was performed with numpy (v.1.24.3) [35] and pandas (v.2.0.3) [36, 37].

Menarche Classification: For the classification of pre- and post-menarche individuals in the ABCD sample, a linear discriminant analysis (LDA) classification model was trained using scikit-learn (version 1.3.0; [38]). For classification we split the full ABCD sample (N=3248 females) into a training and an independent test set by randomly sampling 20% of the data into the test set (visual inspection indicated no difference in menarche status distributions across the whole, training, and test data set; see Additional file 1: Figure S1). The test sample consisted of N=650 participants (pre-menarche: N=419, mean age=11.68 years, $SD=0.56$; post-menarche: N=231, mean age=12.32 years, $SD=0.59$; Table 1). Furthermore, to avoid bias in the training process, propensity score matching [39] was performed in the training data to achieve equal distributions of age and MRI scanner in the pre- and post-menarche groups, as well as equal group sizes. After age-matching, there was no statistical age difference (two-sided independent samples t-test, $p=0.968$) between the pre- and post-menarche groups in the training dataset (pre-menarche: N=775, mean age=12.09 years, $SD=0.58$; post-menarche: N=775, mean age=12.09 years, $SD=0.58$; Table 1).

Participants' responses to the question 'Have you begun to menstruate (started to have your period)?' from the ABCD Youth Pubertal Development Scale and Menstrual Cycle Survey History (PDMS) were used as target labels for the classification algorithm. Responses were encoded numerically in the original survey as (4: Yes; 1: No) and relabelled to 1 and 0.

The estimated class probabilities of the withheld test sample were extracted from the LDA model to further assess the biological validity of the classification. Train- and test set features were independently transformed into z-scores. We implemented a nested cross-validation procedure with a stratified 10-fold outer and inner loop.

Table 1 Demographic information about training and test dataset

	Training data		Test data	
	N	Age _{mean} (std)	N	Age _{mean} (std)
Pre-menarche	775	12.09 (0.58)	419	11.68 (0.56)
Post-menarche	775	12.09 (0.58)	231	12.32 (0.59)

std standard deviation

The inner loop was deployed for hyperparameter tuning via scikit-learn's GridSearchCV. The hyperparameters explored in the grid search included the 'solver' parameter with options ['svd', 'lsqr', 'eigen'] and the 'shrinkage' parameter with values [None, auto, 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]. The performance metric used for evaluation was balanced accuracy and the outer loop was employed to assess the model's performance in the training set. For the final model, an LDA model was re-fitted to the entire training dataset using the selected hyperparameters (solver: least squares solution, shrinkage: 0.7). To assess the model's performance on unseen data, the menarche status of participants from a held-out test sample of ABCD subjects was classified using the final model. A balanced accuracy score and a confusion matrix were calculated. To furthermore confirm the classifier's performance to be above chance in unseen data, we deployed a permutation testing with 1000 iterations. In each iteration, we randomly shuffled the target label (i.e. menarche status) in the training sample and fitted the final model to the training features. Those 1000 random models were then used to predict unseen data in the test sample, resulting in a null distribution of balanced accuracies against which we tested the empirical balanced accuracy score. In detail, we assessed how often the random models would result in a performance metric as extreme as the observed empirical value and divided it by the total number of iterations to obtain a p-value. To foster our results and to control for the influence of the unbalanced test set, we performed an additional analysis in which we iteratively subsampled $n=231$ females 1000 times from the pre-menarche- to match the post-menarche group, which resulted in an evenly balanced test set.

Brain age prediction: The python package of the XGBoost (v2.0.3) library was used [40] to predict chronological age in months from the same 234 sMRI features as those used in the menarche classification. Model tuning was again performed via scikit-learn's GridSearchCV. The hyperparameters explored in the grid search were 'max_depth': [3,6,9], 'max_leaves': [0,2,5,10], 'learning_rate': [0.001,0.01,0.1,0.5,1,3], 'min_child_weight': [1,10,100] and 'n_estimators': [100, 500, 1000]. The performance metric used for evaluation was $r2_score$ and a 5-fold cross-validation approach was employed to assess the model's performance during hyperparameter tuning. The final model was fitted to the entire training dataset using the determined hyperparameters ('learning_rate': 0.01, 'max_depth': 6, 'max_leaves': 0, 'min_child_weight': 10, 'subsample': 0.5, 'num_rounds': 1000). The model's performance on unseen data was tested by applying it to the ABCD withheld test sample described above. The root mean squared error (RMSE) and mean absolute error (MAE) were calculated and the brain age gap

(difference between predicted brain age and chronological age; BAG) was calculated for further analysis. To confirm that the age prediction model performed above chance, a permutation test with 1000 permutations was performed equivalent to the one described above in the menarche classification section.

Association analyses

Association analyses were performed with the ordinary least squares (OLS) regression function of the Python module *statsmodels.formula.api* (v0.14.0) [41]. In line with previous studies [42, 43], a residualised BAG was produced by regressing age and scanning site on BAG. Menarche class probabilities were residualised in the same way to account for age and scanning site. OLS regression was performed to test the association of residualised BAG and residualised menarche class probabilities. Finally, we tested for associations between age at menarche and menarche class probabilities, as well as age at menarche and BAG, controlling for age and scanning site in both instances, using OLS. Likewise, we tested for association between pubertal status, both caregiver- and youth-reported, and menarche class probabilities, as well as BAG, respectively. All association analyses were repeated accounting for potential effects of sociodemographic status (SES), body mass index (BMI) and race / ethnicity. In brief, BMI was calculated by averaging two height and weight measurements respectively and using the formula 'height (lb) / height (in) × 703'. Ethnicity was encoded in 5 levels: White, Black, Hispanic, Asian, other (multiracial or ethnicity with too few members in the sample). Ordinal SES variables were transformed through

rank-based inverse normal transformation and averaged, producing a single SES variable. Further details on the covariates and its calculations can be found in Kraft et al. [13].

Results

We first tested if it was possible to classify from anatomical MRI between same-aged pre- and post-menarcheal girls. Our classifier, trained in a sample of age-matched groups of females pre- and post-menarche performed with a balanced accuracy of 59.24% in the nested cross-validation procedure. Applied to a held-out test set of 419 pre- and 231 post-menarcheal girls, the classifier performed equally well (61.05% balanced accuracy, Fig. 1a). Permutation testing indicated significant above chance performance in the test set ($p=0.001$, Fig. 1b). A validation analysis with an evenly matched subsets (50:50 balanced test set) yielded a mean balanced accuracy of 61.24% (range 57.14–64.94% across 1000 iterations) lending credibility to our initial results (see Additional file 1: Figure S2).

Figure 2a depicts the class probability obtained from the pre-/post-menarche classifier for each individual in the independent ABCD test sample. In an association analysis in the post-menarche group, we found an association of derived class probabilities and age at menarche. Specifically, individuals with an earlier menarche tend to be classified as post-menarche with a higher confidence (coef = -0.0793 , $p=0.002$, see Additional file 1: Figure S3 for bootstrapped results), which may suggest biological sensitivity of the class probabilities beyond the binary pre-/post distinction (Fig. 2b).

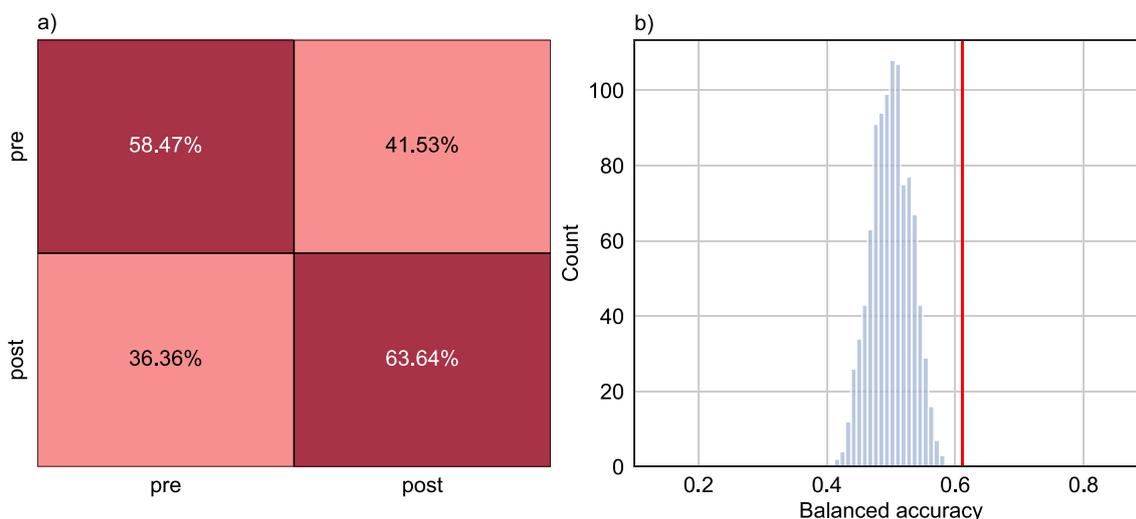


Fig. 1 Menarche can be classified from brain imaging data. **a** Confusion matrix of performance in test data; **b** Histogram of the result of a permutation test in the held-out test sample. The red line shows the empirical balanced accuracy score of 0.61

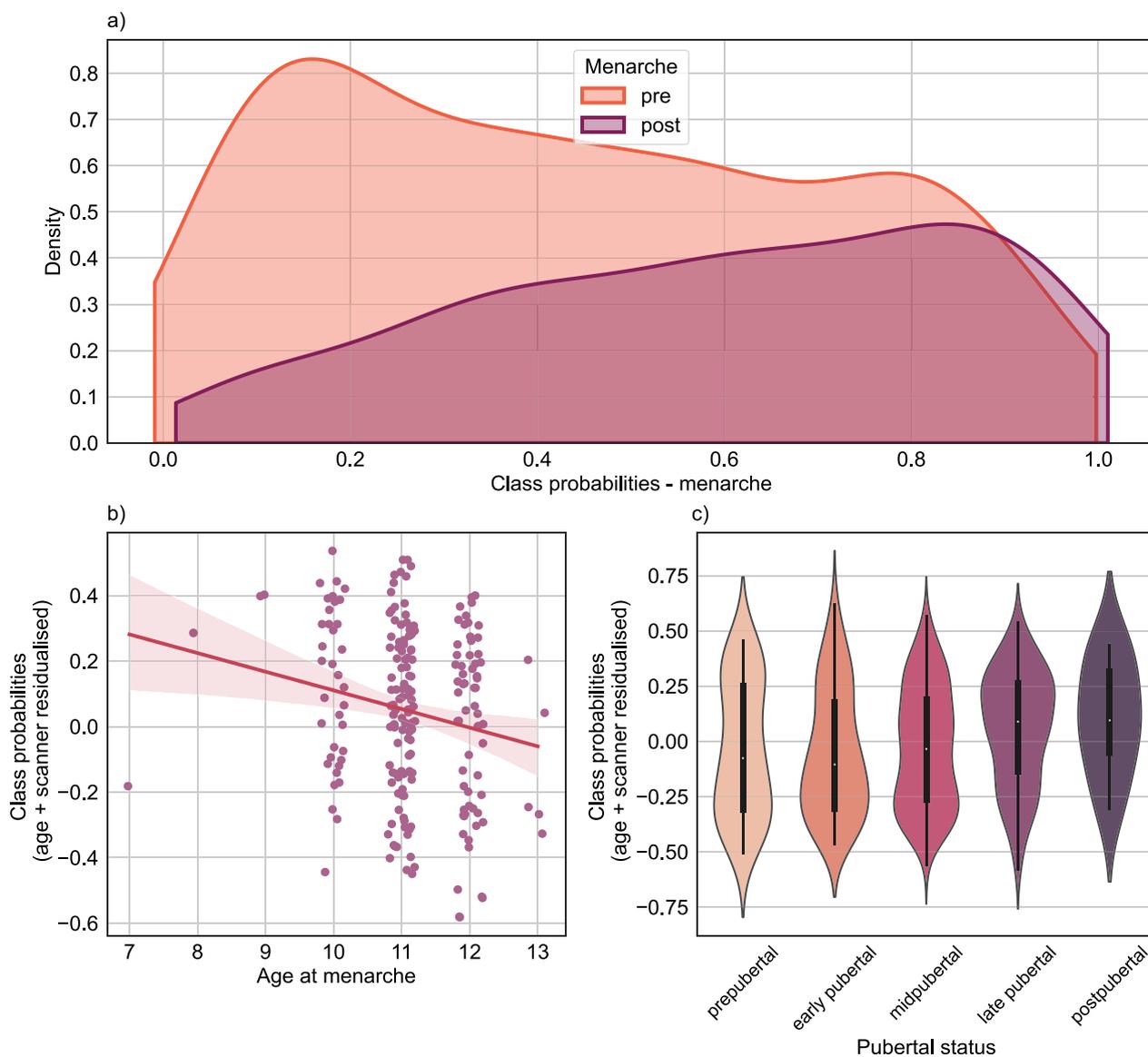


Fig. 2 Menarche class probabilities are associated with measures of pubertal timing and status. **a** Density plot of post-menarche class probabilities of the pre- and post-menarche groups respectively. Class probability of 1 signifies a 100% confident classification as post-menarche, class probability of 0 signifies a 100% confident classification as pre-menarche. **b** Association of age at menarche and menarche class probability controlled for age and scanner. **c** Distribution of class probabilities (age- and scanner residualised) by puberty category (youth-reported)

We furthermore found a positive association between menarche class probability and pubertal status, which was stable across caregiver- and youth-report (youth-reported PDMS: $\text{coef}=0.061$, $p<0.001$, caregiver-reported PDMS: $\text{coef}=0.065$, $p<0.001$; Fig. 2c) supporting our initial interpretation. Of note, the associations with age at menarche and youth-reported puberty status were diminished when incorporating BMI, SES and race/ethnicity as confound factors (Additional file 1: Table S2), highlighting the complex relationship between these factors, puberty and the brain.

Our classifier distinguished between pre- and post-menarcheal females of same age, thereby essentially distinguishing earlier from later pubertal timing relative to age-matched peers. Since pubertal development and age are intertwined, we further sought to investigate whether and to which degree the menarche class probabilities relate to brain age patterns. We next applied the brain age prediction model to the same independent ABCD test sample as used as test sample in the menarche classification. Here, the prediction model performed with an RMSE of 1.3 years and a MAE of 1.1 years (Fig. 3a).

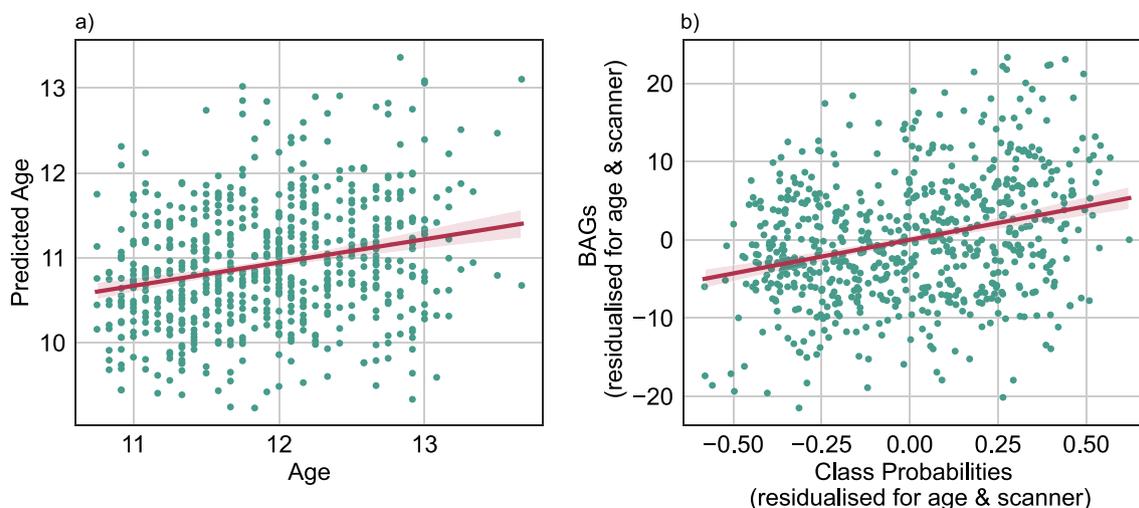


Fig. 3 Comparison of menarche classification to a brain age prediction framework. **a** Predicted age by age. **b** BAGs residualised for age and scanner by menarche class probabilities residualised for age and scanner

Permutation testing indicated performance significantly above chance ($p=0.001$, Additional file 1: Figure S4). From the predicted brain ages, we calculated the brain age gap (difference between predicted brain age and chronological age; BAG). These gaps were significantly associated with menarche class probabilities (Fig. 3b), as observed from a linear model controlling for the effect of age and scanner (coef=8.579, $p<0.001$). This association stayed significant when including BMI, SES, and race / ethnicity as covariates in the analysis (Additional file 1: Table S2). BAG was positively associated with pubertal status (youth-reported PDMS: coef=0.959, $p=0.025$, caregiver-reported PDMS: coef=0.961, $p=0.027$). This effect was descriptively smaller as compared to the class probability effect and also diminished after including both variables into a single model, in which only class probabilities remained significantly associated with pubertal status (PMDS caregiver: coef=0.4407, $p=0.001$, PMDS youth: coef=0.4268, $p=0.001$). Interestingly, whereas the menarche class probability was weakly associated with age at menarche as reported above, the brain age gaps were not. A model controlling for the effects of age and scanner showed no significant correlation of BAG and age at menarche (coef=-1.428, $p=0.089$), lending support to the idea that the menarche classification model may pick up putative biological variability additional to that revealed by a brain age model.

Discussion

The present study aimed at investigating whether structural MRI data can be used to correctly classify pre- vs post-menarche status in adolescent females, thus shedding light on the neurodevelopment associated with

pubertal timing. For this, we successfully trained a machine learning model for the classification of pre- and post-menarcheal females in the ABCD cohort while strictly controlling for age-related neurodevelopment through age-matching. To further disentangle age- vs. puberty-related patterns in neurodevelopment, we performed subsequent comparison to a brain age prediction framework that predicts chronological age from MRI, revealing shared and distinct variance in the two machine learning approaches. Finally, we investigated if the class probabilities obtained from menarche classification may render a continuous biological marker of pubertal timing that can add relevant information beyond the pre- vs post-menarche dichotomy. Indeed, our results indicate that the probabilities are significantly associated with other key variables of pubertal maturation, in particular age at menarche and pubertal status.

Menarche classification

We argue that leveraging a multivariate, machine learning model helps to integrate information from a collection of brain regions into a single score, which eventually may overcome the inherent complexity of modelling puberty in a univariate fashion and its accompanying statistical considerations [7, 44–46] (see [44] for a conceptually similar approach of integrating various sources into a single marker representing pubertal timing). By doing so, our menarche classification model performed with moderate yet significantly above chance accuracy during cross-validation and when applied to a withheld test sample. To rule out that this classification solely mimics a separation of a younger vs. older subgroup of females, we performed a strict age matching prior to model training.

Consequently, our results suggest that there is menarche related neuronal variance detectable in structural MRI data. This aligns well with endocrinological trajectories, which are characterized by a substantial, year-long increase in estradiol levels prior to menarche [47] and related findings that estrogens affect neuroplasticity [6, 48, 49].

Validation of derived class probabilities

Given the close relationship between pubertal- and age-related neurodevelopment we contrasted the results of our classification model to outcomes of a brain age prediction framework. This approach aimed at exploring the degree to which our derived class probabilities and brain-age estimations capture similar or distinct patterns of neurodevelopmental variation in the adolescent female brain. Testing the brain age prediction model on the above mentioned held-out test sample from the ABCD cohort, we observed highly significant and accurate model performance comparable to results of previous studies modelling brain age in the ABCD cohort [6]. Our derived menarche class probabilities were positively related to brain age gaps (BAGs, i.e., the difference between someone's brain and chronological age), matching earlier results that associated brain age with pubertal status (e.g., [6]) and pubertal timing (e.g., [44]). Individuals with higher class probabilities (i.e., a higher probability of being classified as post-menarche) also had higher brain-age gaps (i.e., an indication of a more mature brain in relation to their chronological ages), suggesting that both approaches capture variations in adolescent brain development related to advanced brain maturation. Our work however extends previous findings, by showing that our menarche classification approach seems to be able to better exploit traces of pubertal timing in the brain that are specific to puberty and go beyond the traces of age-related neurodevelopment that are captured by a brain age prediction framework. This finding is in line with previous suggestions that puberty related processes may be a more sensitive measure to investigate adolescent brain development compared to age-related neurodevelopment [6–9].

In particular, we aimed at showing that exploiting the obtained class-probabilities beyond a binary <or>0.5 decision could help towards developing an objective brain-based marker for pubertal development. To prove the additional benefit of such an approach we aimed at investigating the probabilities' associations with puberty-related measurements. In the ABCD study, puberty is assessed by different means, ranging from hormonal measurements to self-reported evaluation of perceived pubertal maturation (see [50]). The latter allows to localize individuals in different pubertal stages or categories

ranging from pre- to post-pubertal. As described before, for females the score is derived by summing over ratings of key physical changes, such as breast development and pubic hair growth, but also the (non-) completion of menarche [14]. Since menarche is directly incorporated in the pubertal category scores, we additionally showed that higher pubertal category scores (thus indicating a later pubertal stage) are associated with higher class probabilities, which serves as an important sanity check for our proposed approach. Higher BAGs were also significantly associated with higher pubertal category scores, however with a descriptively smaller effect sized compared to the class probabilities. Furthermore, after including both variables into a model, only the effect of class probabilities remained significant. Furthermore, we show that post-menarche class probabilities were weakly yet significantly associated with age at menarche. In contrast, the association between BAG and age at menarche was not significant. This may suggest that there are traces of pubertal timing in the brain that go beyond patterns of age-related brain development, and that these traces can be more successfully exploited by our proposed menarche classification model than by the brain age prediction framework. Interestingly, the pattern of higher class probabilities in females that underwent early or earlier menarche resonates with the hypothesis that the brain might be more susceptible to the hormonal influences of puberty at a younger age and that, therefore, individuals who experience an earlier menarche undergo the increase of gonadal hormones at a time when their brain is relatively more sensitive to their effects on neuroplasticity [51, 52]. Of note, after adding SES, race/ethnicity, and BMI as covariates into our analyses, the association between the class probabilities and age at menarche diminished. This observation matches previous reports about the close link between pubertal processes, for example pubertal timing (e.g., operationalized as age at menarche) [53] and these covariates (e.g., [54–56]). Importantly, associations between the aforementioned covariates and puberty were also replicated for ABCD 2-year follow up data, which we used in the current study [14]. While we consider it important to understand the associations with these covariates, their interplay is difficult to disentangle with the data at hand, given the relationship between pubertal timing and these variables. We argue that these results rather warrant further systematic investigation of the interplay between all factors in the equation.

Methodological considerations and future directions

Potential limitations may stem from the fact that we limited our machine learning model to structural imaging data from cortical and subcortical regions. While this

decision resonates with well-replicated findings of cortical and subcortical grey matter changes during puberty [7], integrating additional imaging features, such as white matter measures, may help in a more holistic investigation of pubertal timing. While myelination plays a crucial role in shaping the human brain during adolescence, findings regarding pubertal maturation appear to be either mixed regarding different measures of white matter (e.g., [5, 7]) or lack previous investigations. Furthermore, while we followed a common approach of training our brain-age prediction model in an independent dataset (see e.g., [6, 42]) and applying it to our target sample in the ABCD cohort, recent work from Ray and colleagues [57] suggests that refined brain age models (i.e. a combination of pre-trained models with subsequent finetuning on a fraction of the target data) may improve model performance and thus also downstream analyses. Furthermore, our work focuses on a proof of concept on the 2-year follow up data of the ABCD study. With additional longitudinal data becoming available through upcoming releases, the ABCD study depicts an unprecedented resource to validate our model and proof its usability, since more and more females will eventually undergo their menarche. Lastly, one needs to acknowledge that the validation analyses with class-probabilities and external measures of pubertal timing yielded small—yet significant—effects. Further research is thus warranted to replicate our findings in a different cohort, or later timepoint in the ABCD study, to further elucidate the usability of the continuous class probabilities as proposed here.

Perspective and significance

This work may be seen as a proof of principle that pubertal timing can be classified from brain imaging data. Previous studies that have used age-focused approaches like brain age prediction frameworks have found associations with pubertal measures [44], yet our results suggest that grounding the modelling in puberty data directly may yield brain based markers that are even more sensitive to pubertal status and timing. Future studies may thus further explore similar approaches toward the development of brain-based puberty markers that may be useful in downstream analyses in developmental neuroscience.

Conclusion

We introduced a machine learning approach that classifies menarche status of adolescent females from their cortical and subcortical structural MRI data. The derived continuous brain-based class probabilities captured shared but also unique variations of adolescent neurodevelopment when compared to a brain-age prediction model. Taken together, our results suggest that there are markers of menarche in the brain that can be formalized

into a continuous class probability, which might in the future be developed toward an objective brain-based marker of pubertal timing.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13293-024-00604-4>.

Additional file 1. Methods; Tables S1, S2; Figures S1, S2, S3, S4.

Acknowledgements

The authors used data from the Philadelphia Neurodevelopmental Cohort (PNC, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v3.p2, access permission no 29782) and the Adolescent Brain Cognitive DevelopmentSM Study (ABCD, abcdstudy.org). Support for the collection of the PNC data set was provided by Grant #RC2MH089983 awarded to Raquel Gur, MD, PhD, and #RC2MH089924 awarded to Hakon Hakonarson, MD, PhD. ABCD data, held in the NIMH Data Archive (NDA), is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study[®] is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at <https://abcdstudy.org/federal-partners.html>. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. PNC and ABCD consortium investigators designed and implemented the respective studies and/or provided data but did not participate in the analysis or writing of this report. This manuscript reflects the views of the authors and does not necessarily reflect the opinions or views of any other agency, organization, employer or company.

Author contributions

NG had the initial idea to build a menarche classifier. NG and TK conceptualized the study. TK processed the imaging data. DK curated the ABCD data. NG and DK analysed the data. NG visualized the data. All authors interpreted the data and wrote and revised the paper. TK received funding.

Funding

Open Access funding enabled and organized by Projekt DEAL. TK received support by the Fortüne Program (2660-0-0), Faculty of Medicine, University of Tübingen, the Research Council of Norway (#323961), and the European Research Council (ERC CoG, #101086793, HealthyMom). TK is a member of the Machine Learning Cluster of Excellence, EXC number 2064/1, Project number 39072764. Brain imaging data was processed on the BMBF-funded de.NBI Cloud as part of the German Network for Bioinformatics Infrastructure (031A537B, 031A533A, 031A538A, 031A533B, 031A535A, 031A537C, 031A534A, 031A532B). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the funder. Neither the European Union nor the granting authority can be held responsible for them.

Availability of data and materials

Data used in this study has been shared with the authors under the data usage agreements of the respective cohorts (i.e., ABCD and PNC). To use these resources, interested scholars can obtain access by signing their own respective data usage agreements with the officials. Code used to analyze the data will be made available on github (https://github.com/NGottschewsky/menarche_classification).

Declarations

Ethics approval and consent to participate

PNC study procedures were approved by institutional review boards of the University of Pennsylvania and the Children's Hospital of Philadelphia. All

participants or their caregiver provided written informed consent. ABCD procedures were approved by either the local site Institutional Review Board or by local Institutional Review Board reliance agreements with the central Institutional Review Board at the University of California. All participants and their parents provided written informed consent.

Consent for publication

All subjects consented to study inclusion as part of the data acquisition in the ABCD or PNC cohort. Data was shared with the authors via dedicated data use agreements.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Psychiatry and Psychotherapy, Tübingen Center for Mental Health, University of Tübingen, Tübingen, Germany. ²Centre for Precision Psychiatry, Division of Mental Health and Addiction, Institute of Clinical Medicine, University of Oslo, Oslo, Norway. ³German Center for Mental Health (DZPG), Partner Site Tübingen, Tübingen, Germany. ⁴Max Planck School of Cognition, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.

Received: 31 August 2023 Accepted: 18 March 2024

Published online: 26 March 2024

References

- Mendle J, Beltz AM, Carter R, Dorn LD. Understanding puberty and its measurement: ideas for research in a new generation. *J Res Adolesc.* 2019;29:82–95. <https://doi.org/10.1111/jora.12371>.
- Susman EJ, Dorn LD. Puberty: its role in development. *Handb. Psychol. Dev. Psychol.* Vol 6 2nd Ed, Hoboken: John Wiley & Sons, Inc.; 2013, p. 289–320.
- Sisk CL, Zehr JL. Pubertal hormones organize the adolescent brain and behavior. *Front Neuroendocrinol.* 2005;26:163–74. <https://doi.org/10.1016/j.yfrne.2005.10.003>.
- Paus T, Keshavan M, Giedd JN. Why do many psychiatric disorders emerge during adolescence? *Nat Rev Neurosci.* 2008;9:947–57. <https://doi.org/10.1038/nrn2513>.
- Beck D, Ferschmann L, MacSweeney N, Norbom LB, Wiker T, Aksnes E, et al. Puberty differentially predicts brain maturation in male and female youth: a longitudinal ABCD Study 2022. <https://doi.org/10.1101/2022.12.22.2283852>.
- Holm MC, Leonardsen EH, Beck D, Dahl A, Kjelkenes R, de Lange A-MG, et al. Linking brain maturation and puberty during early adolescence using longitudinal brain age prediction in the ABCD cohort. *Dev Cogn Neurosci.* 2023;60:101220. <https://doi.org/10.1016/j.dcn.2023.101220>.
- Vijayakumar N, Op de Macks Z, Shirtcliff EA, Pfeifer JH. Puberty and the human brain: insights into adolescent development. *Neurosci Biobehav Rev.* 2018;92:417–36. <https://doi.org/10.1016/j.neubiorev.2018.06.004>.
- Blakemore S-J, Burnett S, Dahl RE. The role of puberty in the developing adolescent brain. *Hum Brain Mapp.* 2010;31:926–33. <https://doi.org/10.1002/hbm.21052>.
- Suleiman AB, Galván A, Harden KP, Dahl RE. Becoming a sexual being: the 'elephant in the room' of adolescent brain development. *Dev Cogn Neurosci.* 2017;25:209–20. <https://doi.org/10.1016/j.dcn.2016.09.004>.
- Koolschijn PCMP, Peper JS, Crone EA. The influence of sex steroids on structural brain maturation in adolescence. *PLoS ONE.* 2014;9: e83929. <https://doi.org/10.1371/journal.pone.0083929>.
- Satterthwaite TD, Vandekar S, Wolf DH, Ruparel K, Roalf DR, Jackson C, et al. Sex differences in the effect of puberty on hippocampal morphology. *J Am Acad Child Adolesc Psychiatry.* 2014;53:341–350.e1. <https://doi.org/10.1016/j.jaac.2013.12.002>.
- Wierenga LM, Sexton JA, Laake P, Giedd JN, Tamnes CK, the Pediatric Imaging N and Genetics Study. A key characteristic of sex differences in the developing brain: greater variability in brain structure of boys than girls. *Cereb Cortex.* 2018;28:2741–51. <https://doi.org/10.1093/cercor/bhx154>.
- Kraft D, Alnæs D, Kaufmann T. Traces of pubertal brain development and health revealed through domain adapted brain network fusion. medRxiv; 2023. <https://doi.org/10.1101/2023.01.26.23285055>.
- Herting MM, Uban KA, Gonzalez MR, Baker FC, Kan EC, Thompson WK, et al. Correspondence between perceived pubertal development and hormone levels in 9–10 year-olds from the adolescent brain cognitive development study. *Front Endocrinol.* 2021;11: 549928. <https://doi.org/10.3389/fendo.2020.549928>.
- Savin-Williams RC, Ream GL. Pubertal onset and sexual orientation in an adolescent national probability sample. *Arch Sex Behav.* 2006;35:279–86. <https://doi.org/10.1007/s10508-006-9032-8>.
- Graber JA, Seeley JR, Brooks-Gunn J, Lewinsohn PM. Is pubertal timing associated with psychopathology in young adulthood? *J Am Acad Child Adolesc Psychiatry.* 2004;43:718–26. <https://doi.org/10.1097/01.chi.0000120022.14101.11>.
- Silventoinen K, Haukka J, Dunkel L, Tynelius P, Rasmussen F. Genetics of pubertal timing and its associations with relative weight in childhood and adult height: the swedish young male twins study. *Pediatrics.* 2008;121:e885–91. <https://doi.org/10.1542/peds.2007-1615>.
- Wohlfahrt-Veje C, Mouritsen A, Hagen CP, Tinggaard J, Mieritz MG, Boas M, et al. Pubertal onset in boys and girls is influenced by pubertal timing of both parents. *J Clin Endocrinol Metab.* 2016;101:2667–74. <https://doi.org/10.1210/jc.2016-1073>.
- Ahmed ML, Ong KK, Dunger DB. Childhood obesity and the timing of puberty. *Trends Endocrinol Metab.* 2009;20:237–42. <https://doi.org/10.1016/j.tem.2009.02.004>.
- Copeland WE, Worthman C, Shanahan L, Costello EJ, Angold A. Early pubertal timing and testosterone associated with higher levels of adolescent depression in girls. *J Am Acad Child Adolesc Psychiatry.* 2019;58:1197–206. <https://doi.org/10.1016/j.jaac.2019.02.007>.
- Mendle J, Ferrero J. Detrimental psychological outcomes associated with pubertal timing in adolescent boys. *Dev Rev.* 2012;32:49–66. <https://doi.org/10.1016/j.dr.2011.11.001>.
- Stumper A, Olino TM, Abramson LY, Alloy LB. Pubertal timing and substance use in adolescence: an investigation of two cognitive moderators. *J Abnorm Child Psychol.* 2019;47:1509–20. <https://doi.org/10.1007/s10802-019-00524-7>.
- Hoyt LT, Niu L, Pachucki MC, Chaku N. Timing of puberty in boys and girls: implications for population health. *SSM Popul Health.* 2020;10: 100549. <https://doi.org/10.1016/j.ssmph.2020.100549>.
- McNeilly EA, Saragosa-Harris NM, Mills KL, Dahl RE, Magis-Weinberg L. Reward sensitivity and internalizing symptoms during the transition to puberty: an examination of 9-and 10-year-olds in the ABCD Study. *Dev Cogn Neurosci.* 2022;58: 101172. <https://doi.org/10.1016/j.dcn.2022.101172>.
- Trépanier L, Juster R-P, Marin M-F, Plusquellec P, Francois N, Sindi S, et al. Early menarche predicts increased depressive symptoms and cortisol levels in Quebec girls ages 11 to 13. *Dev Psychopathol.* 2013;25:1017–27. <https://doi.org/10.1017/S0954579413000345>.
- Volkow ND, Koob GF, Croyle RT, Bianchi DW, Gordon JA, Koroshetz WJ, et al. The conception of the ABCD study: from substance use to a broad NIH collaboration. *Dev Cogn Neurosci.* 2018;32:4–7. <https://doi.org/10.1016/j.dcn.2017.10.002>.
- Hagler DJ, Hatton Sean N, Cornejo MD, Makowski C, Fair DA, Dick AS, et al. Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *Neuroimage.* 2019;202:116091. <https://doi.org/10.1016/j.neuroimage.2019.116091>.
- Petersen AC, Crockett L, Richards M, Boxer A. A self-report measure of pubertal status: reliability, validity, and initial norms. *J Youth Adolesc.* 1988;17:117–33. <https://doi.org/10.1007/BF01537962>.
- Satterthwaite TD, Connolly JJ, Ruparel K, Calkins ME, Jackson C, Elliott MA, et al. The Philadelphia Neurodevelopmental Cohort: a publicly available resource for the study of normal and abnormal brain development in youth. *Neuroimage.* 2016;124:1115–9. <https://doi.org/10.1016/j.neuroimage.2015.03.056>.
- Fischl B. FreeSurfer. *Neuroimage.* 2012;62:774–81. <https://doi.org/10.1016/j.neuroimage.2012.01.021>.

31. Rosen AFG, Roalf DR, Ruparel K, Blake J, Seelaus K, Villa LP, et al. Quantitative assessment of structural image quality. *Neuroimage*. 2018;169:407–18. <https://doi.org/10.1016/j.neuroimage.2017.12.059>.
32. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006;31:968–80. <https://doi.org/10.1016/j.neuroimage.2006.01.021>.
33. Fortin J-P, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*. 2018;167:104–20. <https://doi.org/10.1016/j.neuroimage.2017.11.024>.
34. Van Rossum G, Drake FL. The Python language reference. Release 3.0.1 [Repr.]. Hampton, NH: Python Software Foundation; 2010.
35. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. *Nature*. 2020;585:357–62. <https://doi.org/10.1038/s41586-020-2649-2>.
36. McKinney W. Data structures for statistical computing in Python, Austin, Texas: 2010, p. 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>.
37. Reback J, Jbrockmendl, McKinney W, Van Den Bossche J, Roeschke M, Augspurger T, et al. pandas-dev/pandas: Pandas 1.4.3 2022. <https://doi.org/10.5281/ZENODO.6702671>.
38. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12:2825–30.
39. Zhao Q-Y, Luo J-C, Su Y, Zhang Y-J, Tu G-W, Luo Z. Propensity score matching with R: conventional methods and new features. *Ann Transl Med*. 2021;9:812. <https://doi.org/10.21037/atm-20-3998>.
40. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., New York, NY, USA: Association for Computing Machinery; 2016, p. 785–94. <https://doi.org/10.1145/2939672.2939785>.
41. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with python, Austin, Texas: 2010, p. 92–6. <https://doi.org/10.25080/Majora-92bf1922-011>.
42. Kaufmann T, van der Meer D, Doan NT, Schwarz E, Lund MJ, Agartz I, et al. Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nat Neurosci*. 2019;22:1617–23. <https://doi.org/10.1038/s41593-019-0471-7>.
43. Trang TL, Kuplicki RT, McKinney BA, Yeh H-W, Thompson WK, Paulus MP, et al. A nonlinear simulation framework supports adjusting for age when analyzing BrainAGE. *Front Aging Neurosci*. 2018;10.
44. Dehestani N, Whittle S, Vijayakumar N, Silk TJ. Developmental brain changes during puberty and associations with mental health problems. *Dev Cogn Neurosci*. 2023;60: 101227. <https://doi.org/10.1016/j.dcn.2023.101227>.
45. Matte Bon G, Kraft D, Comasco E, Derntl B, Kaufmann T. Modeling brain sex in the limbic system as phenotype for female-prevalent mental disorders. medRxiv; 2023. <https://doi.org/10.1101/2023.08.17.23294165>.
46. Rosenberg MD, Finn ES. How to establish robust brain–behavior relationships without thousands of individuals. *Nat Neurosci*. 2022;25:835–7. <https://doi.org/10.1038/s41593-022-01110-9>.
47. DiVall SA, Radovick S. Pubertal development and Menarche. *Ann N Y Acad Sci*. 2008;1135:19–28. <https://doi.org/10.1196/annals.1429.026>.
48. Brinton RD. Estrogen-induced plasticity from cells to circuits: predictions for cognitive function. *Trends Pharmacol Sci*. 2009;30:212–22. <https://doi.org/10.1016/j.tips.2008.12.006>.
49. Galea LAM, Frick KM, Hampson E, Sohrabji F, Choleris E. Why estrogens matter for behavior and brain health. *Neurosci Biobehav Rev*. 2017;76:363–79. <https://doi.org/10.1016/j.neubiorev.2016.03.024>.
50. Cheng TW, Magis-Weinberg L, Guazzelli Williamson V, Ladouceur CD, Whittle SL, Herting MM, et al. A Researcher’s Guide to the measurement and modeling of puberty in the ABCD Study® at baseline. *Front Endocrinol*. 2021;12: 608575. <https://doi.org/10.3389/fendo.2021.608575>.
51. Graber JA. Pubertal timing and the development of psychopathology in adolescence and beyond. *Horm Behav*. 2013;64:262–9. <https://doi.org/10.1016/j.yhbeh.2013.04.003>.
52. Schulz KM, Molenda-Figueira HA, Sisk CL. Back to the future: the organizational–activational hypothesis adapted to puberty and adolescence. *Horm Behav*. 2009;55:597–604. <https://doi.org/10.1016/j.yhbeh.2009.03.010>.
53. Barendse M, Byrne ML, Flournoy J, McNeilly EA, Williamson VG, Barrett A-M, et al. Multi-method assessment of pubertal timing and associations with internalizing psychopathology in early adolescent girls. 2020. <https://doi.org/10.31234/osf.io/p5vfb>.
54. Hiatt RA, Stewart SL, Deardorff J, Danial E, Abdiwahab E, Pinney SM, et al. Childhood socioeconomic status and menarche: a prospective study. *J Adolesc Health*. 2021;69:33–40. <https://doi.org/10.1016/j.jadohealth.2021.02.003>.
55. Kelly Y, Zilanawala A, Sacker A, Hiatt R, Viner R. Early puberty in 11-year-old girls: Millennium Cohort Study findings. *Arch Dis Child*. 2017;102:232–7. <https://doi.org/10.1136/archdischild-2016-310475>.
56. Rosenfield RL, Lipton RB, Drum ML. Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index. *Pediatrics*. 2009;123:84–8. <https://doi.org/10.1542/peds.2008-0146>.
57. Ray B, Chen J, Fu Z, Suresh P, Thapaliya B, Farahdel B, et al. Replication and refinement of brain age model for adolescent development 2023. <https://doi.org/10.1101/2023.08.16.553472>.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.