

1 “Puberty age gap”: New method of assessing pubertal timing and its association with 2 mental health problems

3 Niousha Dehestani^{1,2*}, Nandita Vijayakumar¹, Gareth Ball^{3,4}, Sina Mansour L.^{2,5}, Sarah Whittle^{2,6} &
4 Timothy J. Silk^{1,3,6*}

5

6 ¹School of Psychology, Deakin University, 221 Burwood Highway, Burwood, VIC 3125, Australia.

7 ²Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne and
8 Melbourne Health, Victoria, Australia

9 ³Developmental Imaging, Murdoch Children’s Research Institute, Parkville 3052, Australia.

10 ⁴Department of Paediatrics, The University of Melbourne, Melbourne 3053, Australia.

11 ⁵Department of Biomedical Engineering, The University of Melbourne, Victoria, Australia

12 ⁶These authors jointly supervised this work: Sarah Whittle and Timothy J. Silk

13

14 Correspondence

15 Niousha Dehestani, School of Psychology, Deakin University, 221 Burwood Highway,
16 Burwood, VIC 3125, Australia, ndehestanikolag@deakin.edu.au, +61422937273 or Tim Silk,
17 School of Psychology, Deakin University, 221 Burwood Highway, Burwood, VIC 3125,
18 Australia, tim.silk@deakin.edu.au, +61408533160.

19

“Puberty Age gap”: A new method of pubertal timing

20 Abstract

21

22 Puberty is linked to mental health problems during adolescence, and in particular, the timing
 23 of puberty is thought to be an important risk factor. This study developed a new measure of
 24 pubertal timing using multiple pubertal features and nonlinear associations with age and
 25 investigated its association with mental health problems. Using the Adolescent Brain
 26 Cognitive Development (ABCD) cohort, we implemented three models of pubertal timing by
 27 predicting chronological age from i) observed physical development, ii) hormones
 28 (testosterone and dehydroepiandrosterone [DHEA]), and iii) a combination of the two, using
 29 a supervised machine learning method (n ~9,900). Accuracy of the new models, and their
 30 associations with mental health problems were evaluated. The new pubertal timing measure
 31 performed better in capturing age variance compared to a commonly used method, and the
 32 physical measure accounted for more variance in mental health, such that earlier pubertal
 33 timing was associated with higher symptoms. This study demonstrates the utility for a new
 34 model of pubertal timing and suggests that physical maturation may play a predominant role
 35 in predicting mental health problems in early adolescence.

36

“Puberty Age gap”: A new method of pubertal timing

37 Introduction

38

39 Adolescence is the period between childhood and adulthood when individuals acquire the
40 emotional and cognitive skills necessary to gain independence from their parents(1,2).
41 Individuals also progress through puberty, the process of attaining reproductive maturity,
42 during this period. This includes hormonal and physical changes such as body hair growth
43 and gonadal maturation (along with the development of breasts and onset of menstruation in
44 females)(3,4). The progression through puberty has been associated with an increase in
45 susceptibility to a range of internalizing and externalizing mental health problems(5,6), which
46 may reflect the effects of hormones on the central nervous system and/or psychosocial
47 mechanisms related to physical differences from peers(7). Therefore, investigation of
48 individual differences in pubertal processes that consider both hormonal and physical
49 changes may lead to a better understanding of adolescent mental health problems.

50

51 While all typically developing individuals progress through the same stages of puberty (based
52 on observable physical changes), the onset and speed of progression can differ across
53 individuals. As such, at any given period during adolescence, there is marked variability in
54 pubertal stage, termed “pubertal timing” (8,9). Importantly, it is pubertal timing – not
55 pubertal stage – that is often linked to the emergence and severity of mental health problems
56 (10), though there are inconsistencies in the literature. A number of studies have found that in
57 females, earlier pubertal timing is associated with internalizing problems (11,12), including
58 depression(13,14), anxiety(15), eating disorders (16), and externalizing behaviors. On the
59 other hand, a number of other studies report null findings regarding internalizing
60 symptoms(14,15). Similarly in males, both early and late timing has been related to
61 internalizing (11 VS 16) and externalizing problems (17 VS 18), while others have failed to

“Puberty Age gap”: A new method of pubertal timing

62 identify any associations (9, 15). Small sample sizes and methodological differences between
63 studies may partially explain the inconsistent findings, with variation in the measures used to
64 calculate pubertal timing likely important. Indeed, a meta-analysis of 101 studies found that
65 the method of measuring puberty had a moderating role in the relationship between pubertal
66 timing and mental health problems (4).

67

68 Different statistical approaches exist to measure the relative pubertal timing of
69 individuals compared to same-aged peers. A common approach is to regress age from
70 pubertal status (e.g., based on Pubertal Developmental Scale [PDS] scores) in a sample, with
71 residuals reflecting earlier or later maturity relative to the group average (21). However, such
72 approaches only capture individual differences in observable physical development and do
73 not inform us about underlying biological mechanisms that are more reflected in hormone
74 levels (22). Hormone levels provide valuable information regarding the endocrine processes
75 of puberty (23,24). Previous studies have also demonstrated associations between changes in
76 hormone levels and adolescent mental health problems. For example, increasing levels of
77 dehydroepiandrosterone (DHEA) have been linked to internalizing symptoms (25), while
78 testosterone has been associated with externalizing and disruptive behaviors (26). To the best
79 of our knowledge, only one study has calculated pubertal timing with hormone data to predict
80 internalizing behaviors in a sample of 174 females (27). Although this study did not find any
81 significant associations, it is important to investigate this in a larger sample, of both females
82 and males, and to assess associations across different dimensions of mental health problems.
83 Additionally, no previous studies have combined both hormonal and physical measurements
84 when calculating pubertal timing, which may be more sensitive to multiple underlying
85 mechanisms that contribute to mental health problems. Thus, further investigation is needed

“Puberty Age gap”: A new method of pubertal timing

86 to develop pubertal timing methods that amalgamate different aspects of puberty in a single
87 model.

88

89 In this study, we propose a multivariate method to calculate pubertal timing that
90 draws upon the “brain age” approach (28,29), where an association between multiple
91 neuroimaging variables and chronological age is learned with supervised machine learning
92 methods. Subtracting chronological age from brain age yields a “brain age gap” that reflects
93 brain maturation relative to the group average (26,28). The major benefits of this model
94 include being able to combine multiple features and reducing complex multivariate
95 information to a single parameter. We propose to use a similar model to calculate estimates of
96 pubertal timing by combining multiple puberty-related features. In contrast, alternative
97 strategies either focus on single measures and only capture specific aspects or mechanisms of
98 puberty or calculate the mean of multiple measures but may obscure the relative importance
99 of each item. The second benefit of this method is that it can model nonlinear relationships
100 between multiple measures of puberty and age, which is important given that nonlinear
101 associations between specific features of puberty and chronological age have been observed
102 in previous studies (30,31).

103

104 In the current study, we aimed to create a normative model of pubertal timing that combines
105 multiple measures of puberty capturing hormone levels and physical changes. By using a
106 strict cross-validation approach, we ensure that model performance is robust. We compared
107 model performance to a common linear model of pubertal timing (i.e., regressing age from
108 total PDS score). We also compared the combined normative model against two unimodal
109 models employing a similar supervised machine learning approach, one that uses hormonal
110 measures alone and another that uses physical measures only. Finally, we examined the

“Puberty Age gap”: A new method of pubertal timing

association between each pubertal timing model and multiple dimensions of mental health problems. We hypothesized that the combined normative model of pubertal timing, which uses both hormonal and physical measures, would provide a more accurate prediction of age and better prediction of mental health problems compared to either of the unimodal (hormonal or physical) normative models or the (traditional) linear model of puberty. We additionally hypothesized that early pubertal timing would be associated with increased mental health problems in females and males. Based on a recent meta-analysis, we hypothesized significant effects of similar magnitude for internalizing and externalizing problems, while no significant effects were expected for attention problems (4). Prior findings of sex differences are mixed, and as such we did not propose specific hypotheses here.

Materials and Methods

To promote reproducible open research practices, all analyses conducted as part of this article are made publicly available in a git repository (“<https://github.com/Niousha-Dehestani/Puberty-age>”). In addition, more detail about participants, measurements and analysis is provided in the supplementary materials.

Participants

Participants were drawn from the ongoing, longitudinal, Adolescent Brain Cognitive Development (ABCD) Study (<https://abcdstudy.org/>). Data was collected from ~11,500 children at baseline (47% females, age 9-10 years old) from 21 sites across the United States,

“Puberty Age gap”: A new method of pubertal timing

with annual data collection thereafter (see Supplementary Table 1 for demographic information). Data from baseline and 3 annual follow-up waves were used in the current analyses. We excluded participants who had a mismatch between their biological sex (collected per visit with the salivary sample) and their self-reported gender as well as those with missing values in biological sex (See Supplementary Information (SI) Appendix S1 for details of data cleaning procedures). The exact sample sizes utilized in each analysis are reported in detail below.

Measures

Pubertal Development Scale. The Pubertal Development Scale (PDS) measures observable physical signs of puberty. It includes items on height, body hair, and skin change in both sexes, as well as the onset of menarche and breast development for girls, and facial hair and voice changes for boys. Items are rated on a Likert scale from one to four points (“had not begun” to “already complete”). Onset of menarche was a binary variable (yes/no response) that was converted to one for “no” and four for “yes”. The PDS can either be collected via self- or parent-report. However, each of these measures have their own limitations. The parent-report has good correspondence to clinician ratings, though this correspondence is lower in males (32). Conversely, the self-report is less accurate for individuals who are in the lower or upper pubertal stages as they tend to report toward the mid stages. Therefore, some studies have recommended using the parent-report PDS, especially in late childhood or early adolescence (33), and accordingly, the current study utilized this version. In this study we used individual PDS items rather than the average PDS score typically used in prior literature. This allowed us to capture the unique relationship between each item and our

“Puberty Age gap”: A new method of pubertal timing

outcomes of interest, consistent with previous literature that has found that separate PDS items exhibit differential relationships with, e.g., brain structure (34).

Hormones. DHEA and testosterone (TST) levels were measured via salivary hormone samples assayed by Salimetrics. Although estradiol was also measured, it was not used in the current analyses due to it only being available for females and having excessive missingness (n = 780). The hormone data was cleaned based on the protocol that published recently(35), which involved removing the confounding effects of collection time, duration of collection, wake-up time on collection day, having exercised before collection, and caffeine intake (with a linear mixed effect model, see SI for more details). Further information on the reliability and overall quality of the hormone data can be found in protocol (32).

Body Mass Index (BMI). BMI was calculated as the average of two weight and height measurements per visit, assessed by the researcher. Next, BMI standard deviation scores (BMI z-scores) were calculated relative to age and sex, with reference to the CDC 2000 Growth Charts (36).

Sociodemographic Variables. Five categories of race/ethnicity were coded: White, Black, Hispanic, Asian, and Other/Multi-race. Additionally, household income and education were obtained as measures of family socioeconomic status (SES) that were collected in a parent report.

Mental health problems. The Child Behavior Checklist (CBCL) (age 6 to 18 form; (37)) was used to measure parent-reported mental health problems. CBCL includes 8 syndrome scales;

“Puberty Age gap”: A new method of pubertal timing

namely, Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Aggressive Behavior, and Rule-breaking Behavior as well as three broad summary scales of externalizing, internalizing, and total problems.

Statistical Analysis

Calculating pubertal timing. Inheriting fundamental concepts from the literature on “brain age” (26), “puberty age” was computed using supervised machine learning. The model was trained to learn the relationship between physical and hormonal measurements of puberty (specifically, each PDS item, DHEA levels and TST levels) and chronological age, separately in males and females. Thereafter, the model was used to predict chronological age from pubertal measurements in an independent test sample. The effect of age was subsequently regressed from these predictions to adjust the bias created by regression toward the mean (RTM) (For more details see SI, Appendix S3, Figure S2 and Figure S3, and (38)). The bias-adjusted prediction of an individual’s chronological age from pubertal measurements is termed “puberty age”. Further, the residuals of the prediction model (after subtracting chronological age from puberty age) are referred to as the “puberty age gap”, which we use as a dimensional measure indicative of relative pubertal timing. A positive puberty age gap is interpreted as a sign of earlier pubertal development compared to the normative age and sex-matched group, while a negative gap reflects relatively delayed pubertal development (see, Appendix 1, Figure S1).

We implemented a Generalized Additive Model (GAM) for the prediction of age (response variable) from pubertal measurements, which utilized TST, DHEA, and each of the PDS items as multivariate predictor variables. To remove the potential impacts of familial relations and repeat assessments of each individual, the sample was stratified to randomly

“Puberty Age gap”: A new method of pubertal timing

keep a single observation for each family (i.e., across waves and siblings). For details on data processing, such as the procedure for dealing with missing values, see SI Appendix S1 and Table S1. Additionally, to ensure the model was trained on a typically developing sample; using the CBCL DSM-oriented scales (consistent with DSM diagnostic categories), we excluded individuals above the threshold (symptom’s score >60) for affective, anxiety, somatic, opposition defiant and conduct problems, as well as ADHD. The remaining participants who had symptom scores <60 were included in the training of models of puberty age (N = 4,949 (2,439 females)). However, a representation of the whole sample was used for testing (N = 9,919 (4,725 females)).

As GAM fits smooth nonlinear curves in the form of spline functions, it is expected to outperform commonly utilized linear methods of pubertal timing measurement, given the nonlinear relations previously reported(30). Model training was performed within a nested 10-fold cross-validation design (for each fold in the outer loop, 90% of samples were used for model training and 10% for testing). Inner loop hyperparameter tuning was performed by a grid search for optimal regularization penalty on each term (i.e, using GAM) to minimize the estimated prediction error in the training sample (generalized cross-validation (GCV) score). To understand the contribution of different indices to the measurement of pubertal timing, we implemented three alternative models to estimate “puberty age”. The first approach only used hormones (DHEA and TST) as features to predict chronological age, the second only used PDS items, and the third combined hormone and PDS items. The predictions of age were named *hormonal* puberty age, *physical* puberty age, and *combined* puberty age, respectively. In all three models, the residuals of age prediction indexed “puberty age gap”, which reflects pubertal timing. We also used the partial dependence function in GAM models that can

“Puberty Age gap”: A new method of pubertal timing

reflect the importance of each feature in the combined puberty age model (for more detail, see SI, Appendix S4 and Figure S4).

Comparison of “puberty age” models. We compared the performance of the three alternate puberty age models using Pearson's correlation between predicted and chronological age and median absolute error (MAE) averaged over the 10 cross-validation folds. Additionally, we used a non-parametric paired t-test (Wilcoxon signed-rank test) to statistically compare the model performance of different models based on the absolute error of predictions.

“Physical puberty age” compared to a traditional pubertal timing model. The current study also compared the out-of-sample performance and accuracy of the physical puberty age model with the most common traditional pubertal timing model: linearly regressing age from average PDS score. In order to draw a comparison between the new model and the traditional approach, the performance of the physical puberty age model was contrasted with the traditional approach. The physical model was selected (rather than the hormonal or combined model) to ensure a fair comparison of two approaches that measure pubertal timing from the same input features (i.e., PDS). In order to conduct this comparison, a linear regression model was used to regress chronological age from the total PDS score (train) and used the fitted model coefficients in unseen data to measure timing (test). Similar to the puberty age model design, a 10-fold cross-validation design was used to measure traditional pubertal timing for the whole sample. This linear model provided an implementation of the traditional model in an out-of-sample prediction paradigm. Model performance was assessed based on the out-of-sample prediction accuracies (quantified by the absolute error of predictions). Similarly, we used the “Wilcoxon signed-rank test” to investigate the statistical differences in model performance.

“Puberty Age gap”: A new method of pubertal timing

“Puberty age gap” associations with mental health problems. We used linear mixed-effect models (LMM) to investigate associations between each alternate “puberty age gap” measure and different dimensions of mental health problems, in males and females separately. The following formula was tested for each syndrome dimension (i.e., Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Aggressive Behavior, and Rule-breaking Behavior), as well as three broad scales including total problems, externalizing, and internalizing problems. Sample size for this analysis was $N = 9,919$ (4,725 (females)) after removing the missing value in CBCL items.

Mental health problems \sim *puberty age gap* + *age* + (*1/site*).

Age was included as a confound (fixed effect) and data collection sites were modelled as random effects. We corrected for multiple tests controlling for False Discovery Rate (FDR) at 5% and reported the FDR corrected p-values in the results. The different “pubertal age” LMMs were compared based on the Akaike Information Criterion (AIC) and a cut-off of 2 was used to indicate evidence for a better model, i.e., the model with an AIC that is at least 2 units smaller is considered a comparatively better model. Furthermore, to investigate whether associations between pubertal timing and mental health problems differed by age, we investigated the interaction effect of age and puberty age gap in predicting mental health problems using the following LMM:

Mental health problems \sim *puberty age gap* + *age* + *puberty age gap: age* + (*1/site*)

“Puberty Age gap”: A new method of pubertal timing

Finally, given the known association of pubertal timing with BMI, SES, and race/ethnicity (Mendle et al., 2019), supplementary analyses repeated primary models while accounting for these variables as covariates (see SI, Appendix S5 and Table S2, for details of these analyses). This approach avoided potential complexity in our primary analyses due to the collinearity of confounding variables with our main variables of interest.

Results

Accuracy of puberty age models. All three models were able to provide significant and accurate out-of-sample predictions of age (Table 1, Figure 1). The performance of the hormonal model was better than the physical model in males, while on the contrary, physical puberty age was found to be a better predictor of age than hormonal puberty age in females. Across both sexes, combined puberty age explained the largest degree of variation in chronological age. Non-parametric statistical comparisons showed that the MAE between all models was significantly different ($p < 0.001$). See the correlation between these three models in SI (See Appendix S2 and Figure S1).

Approximate position of Table 1

Approximate position of Figure 1

“Puberty age model” vs. Traditional model. Comparing the out-of-sample prediction performance of the physical puberty age gap with the traditional model of pubertal timing showed that the physical puberty age model was significantly better at modelling the

“Puberty Age gap”: A new method of pubertal timing

relationship between puberty development and age in previously unseen data (See Table 2).

Non-parametric statistical comparisons showed that the MAE between models were

significantly different ($p < 0.001$)

Approximate position of Table 2

Pubertal timing measures predicting mental health problems. We investigated the association between each puberty age gap measure (hormonal, physical, and combined) and different dimensions of mental health problems (see Table 3, Figure 2). Hormonal puberty age gap was not significantly associated with any mental health problems in either males or females. Physical puberty age gap, however, was significantly positively associated with all dimensions of mental health problems in males and all dimensions except Anxiety-Depression in females. While similar associations were present for the combined puberty age gap, AIC differences indicated a better model fit for the physical puberty age gap in all cases. Finally, there were no significant interactions between age and any puberty age gap measure in predicting mental health problems in either females or males, suggesting that associations were stable across the age range.

Approximate position of Table 3

Approximate position of Figure 2

Discussion

In this study, we utilized biological and physical pubertal features and used supervised machine learning to model pubertal timing and its association with mental health

“Puberty Age gap”: A new method of pubertal timing

problems. Overall, a combined puberty age model predicted age more accurately than the unimodal (i.e., physical, or hormonal features alone). In males, the performance of the hormonal model was better than the physical model, which was not the case for females. This may be due to the fact that the two hormones tested (DHEA and testosterone) are more predominant in males, whilst female-dominant hormones (such as estradiol) were not considered. We showed that our novel pubertal timing model (based on physical pubertal features) had better accuracy than a conventional linear regression-based measurement in both sexes. Findings also highlight that our new method of assessing pubertal timing was significantly associated with most dimensions of mental health problems in males and females. Generally, we found that early pubertal maturation (based on physical and combined physical and hormonal features) was consistently predictive of various dimensions of mental health problems and could hence be an indicative risk factor for general mental health problems.

Findings indicated that our novel normative puberty age models were able to provide significantly accurate out-of-sample predictions of age. This nonlinear modelling approach was found to be more accurate at predicting age than traditional measures of pubertal timing that rely on linear modelling of the relationship between (average) pubertal stage and age. These results highlight the presence of nonlinear relationships between age and pubertal development, consistent with previous literature reporting benefits of other nonlinear methods for calculating pubertal timing(21,31). Additionally, our examination of the collinearity between each of the PDS items showed that, although the items are significantly correlated, the correlations between them are less than 0.35 (for more details, see supplementary section 2). This suggests that each PDS item carried a significant amount of independent information. Our method thus used this independent information in a multivariate design instead of simplifying this information by averaging items to create a

“Puberty Age gap”: A new method of pubertal timing

single PDS score. Furthermore, our out-of-sample prediction provides an evaluation of the generalizability and replicability of study findings (39). We presented how a model that was trained on 90% of the sample could provide accurate out-of-sample predictions in the remaining 10% of the sample. This provides the opportunity for future research to measure pubertal timing in studies with smaller sample sizes, by taking advantage of this pre-trained normative model. This could potentially alleviate some study biases inherent to small sample sizes.

Beyond its benefits for predicting age, our novel method of assessing pubertal timing was also significantly related to mental health problems in adolescents. Across most models of puberty age, we found that relatively early pubertal timing (i.e., positive puberty age gap) was associated with an increase in most dimensions of mental health problems. This is consistent with the maturation disparity hypothesis, whereby a mismatch between physical development and progression of emotional and cognitive development is purported to increase in those with early pubertal timing, which accounts for difficulties navigating the complexities and challenges of this period and may thus result in greater risk for mental health problems (7, 4). In particular, findings showed that physical puberty age gap outperformed combined and hormonal puberty age gap measures in the prediction of mental health problems in both females and males. Thus, although the prediction accuracy of the combined puberty age model was significantly better than physical puberty age, it was not better at explaining mental health problems. Moreover, we did not find any significant associations between hormonal puberty age gap and mental health problems in males or females. These findings are partially consistent with (40), who reported that pubertal timing measured from hormonal information (testosterone and DHEA) did not predict internalizing symptoms. Consistent with prior studies, these findings may suggest that psychosocial mechanisms have a larger role (in contrast to biological mechanisms indexed by hormones)

“Puberty Age gap”: A new method of pubertal timing

in predicting mental health problems in early adolescence (14). Likewise, prior work has suggested that earlier physical development impacts social functions such as difficulty maintaining friendship with peers who mature later, and a tendency to associate with older adolescents that engage in more externalizing behaviors (12).

While this study has strengths in its large sample size, use of hormonal assays and physical measurement of puberty, as well as implementation of a novel, generalizable method for assessing pubertal timing, there are limitations that should be addressed in future work. First, the age range of the sample used was rather limited and a wider age range could have captured more variance in pubertal maturation and improved the out-of-sample model prediction of age. Future studies that implement our model in a wider age range encompassing the start to completion of puberty would be able to better capture the complete nonlinear relationship between age and pubertal development. Relatedly, while we did not observe any interactive effects of age with pubertal timing in predicting mental health problems symptoms in early adolescents, it is possible that such effects may be detectable across a larger age range. Further, different measures of puberty age gap may also differentially predict mental health problems in earlier versus later adolescence. Second, this study focused on a limited number of pubertal hormones, and the omission of other hormonal features may have impacted the predictive power of hormonal puberty age gap – particularly in females. Additionally, variations in the time for collecting hormones across the day could have also had an impact on the inferior performance of the hormonal age gap (although these were statistically accounted for). Thus, future studies could incorporate more detailed assessments of hormones to improve model prediction. Although the current study used the parent report version of PDS due to reportedly higher reliability in this particular age range, it would be valuable for future research to replicate this analysis in a sample with a wider age range with self- or clinician-assessed pubertal development.

“Puberty Age gap”: A new method of pubertal timing

In conclusion, the current study proposes a smooth nonlinear puberty age model that facilitates generalizable investigations of pubertal timing in future studies. Our findings also highlight the importance of physical pubertal maturation, relative to hormonal changes, for mental health problems during early adolescence. This suggests that psychosocial mechanisms may play an important role in the relationship between early pubertal timing and mental health problems, which has implications for interventions aimed at reducing the risk of the emergence of mental health problems in adolescence.

Data Sharing:

This study used data from Adolescent Brain Cognitive Development (ABCD) Study (<https://abcdstudy.org>), held in the NIMH Data Archive (NDA). This is a longitudinal study which collected data from ~ 11000 children in age 9-10 and follow them through 10 years. 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/scientists/workgroups/>. ABCD consortium investigators designed and implemented the study 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 may not reflect the opinions or views of the NIH or ABCD consortium investigators.

Acknowledgement

Niousha Dehestani was supported by a Deakin University Postgraduate

“Puberty Age gap”: A new method of pubertal timing

425 Research Scholarships (DUPRS). Additionally, this study was supported by the MASSIVE
426 high-performance computing facility (www.massive.org.au).

427

428 **Conflict of Interest**

429 The authors declare that they have no financial interests or personal relationships that could
430 impact the work reported in this paper.

431

432

“Puberty Age gap”: A new method of pubertal timing

433

434 **References**

- 435 1. Blakemore S-J, Choudhury S. Development of the adolescent brain: implications for
436 executive function and social cognition. J Child Psychol Psychiatry [Internet]. 2006
437 Mar;47(3–4):296–312. Available from:
438 <https://onlinelibrary.wiley.com/doi/10.1111/j.1469-7610.2006.01611.x>
- 439 2. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence.
440 Lancet Child Adolesc Heal. 2018;2(3):223–8.
- 441 3. Banerjee I, Clayton P. The Genetic Basis for the Timing of Human Puberty. J
442 Neuroendocrinol [Internet]. 2007 Nov;19(11):831–8. Available from:
443 <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2826.2007.01598.x>
- 444 4. Kohen-Raz R. PHYSIOLOGICAL MATURATION AND MENTAL GROWTH AT
445 PRE-ADOLESCENCE AND PUBERTY. J Child Psychol Psychiatry [Internet]. 1974
446 Jul;15(3):199–213. Available from:
447 <https://onlinelibrary.wiley.com/doi/10.1111/j.1469-7610.1974.tb01245.x>
- 448 5. Ullsperger JM, Nikolas MA. A meta-analytic review of the association between
449 pubertal timing and psychopathology in adolescence: Are there sex differences in risk?
450 Psychol Bull [Internet]. 2017 Sep;143(9):903–38. Available from:
451 <http://doi.apa.org/getdoi.cfm?doi=10.1037/bul0000106>
- 452 6. Natsuaki MN, Klimes-Dougan B, Ge X, Shirtcliff EA, Hastings PD, Zahn-Waxler C.
453 Early pubertal maturation and internalizing problems in adolescence: Sex differences
454 in the role of cortisol reactivity to interpersonal stress. J Clin Child Adolesc Psychol.
455 2009;38(4):513–24.
- 456 7. Ge X, Natsuaki MN. In search of explanations for early pubertal timing effects on
457 developmental psychopathology. Curr Dir Psychol Sci. 2009;18(6):327–31.

“Puberty Age gap”: A new method of pubertal timing

- 458 8. A W, TANNER JM. VARIATIONS IN PATTERN OF PUBERTAL CHANGES IN
459 GIRLS. *Obstet Gynecol Surv* [Internet]. 1970 Jul;25(7):694–5. Available from:
460 <http://journals.lww.com/00006254-197007000-00018>
- 461 9. Petersen AC, Crockett L, Richards M, Boxer A. A self-report measure of pubertal
462 status: Reliability, validity, and initial norms. *J Youth Adolesc* [Internet]. 1988
463 Apr;17(2):117–33. Available from: <http://link.springer.com/10.1007/BF01537962>
- 464 10. Graber JA. Pubertal timing and the development of psychopathology in adolescence
465 and beyond. *Horm Behav* [Internet]. 2013 Jul;64(2):262–9. Available from:
466 <http://dx.doi.org/10.1016/j.yhbeh.2013.04.003>
- 467 11. Crockett LJ, Carlo G, Wolff JM, Hope MO. The role of pubertal timing and
468 temperamental vulnerability in adolescents’ internalizing symptoms. *Dev*
469 *Psychopathol* [Internet]. 2013 May 30;25(2):377–89. Available from:
470 [https://www.cambridge.org/core/product/identifier/S0954579412001125/type/journal_](https://www.cambridge.org/core/product/identifier/S0954579412001125/type/journal_article)
471 [article](https://www.cambridge.org/core/product/identifier/S0954579412001125/type/journal_article)
- 472 12. Marceau K, Neiderhiser JM, Lichtenstein P, Reiss D. Genetic and Environmental
473 Influences on the Association Between Pubertal Maturation and Internalizing
474 Symptoms. *J Youth Adolesc*. 2012;41(9):1111–26.
- 475 13. Hamlat EJ, McCormick KC, Young JF, Hankin BL. Early pubertal timing predicts
476 onset and recurrence of depressive episodes in boys and girls. *J Child Psychol*
477 *Psychiatry Allied Discip*. 2020;61(11):1266–74.
- 478 14. Mendle J, Ferrero J. Detrimental psychological outcomes associated with pubertal
479 timing in adolescent boys. *Dev Rev* [Internet]. 2012;32(1):49–66. Available from:
480 <http://dx.doi.org/10.1016/j.dr.2011.11.001>
- 481 15. Weingarden H, Renshaw KD. Early and late perceived pubertal timing as risk factors
482 for anxiety disorders in adult women. *J Psychiatr Res* [Internet]. 2012;46(11):1524–9.

“Puberty Age gap”: A new method of pubertal timing

- 483 Available from: <http://dx.doi.org/10.1016/j.jpsychires.2012.07.015>
- 484 16. Klump KL. Puberty as a critical risk period for eating disorders: A review of human
485 and animal studies. *Horm Behav* [Internet]. 2013;64(2):399–410. Available from:
486 <http://dx.doi.org/10.1016/j.yhbeh.2013.02.019>
- 487 17. Winer JP, Parent J, Forehand R, Breslend NL. Interactive Effects of Psychosocial
488 Stress and Early Pubertal Timing on Youth Depression and Anxiety: Contextual
489 Amplification in Family and Peer Environments. *J Child Fam Stud* [Internet]. 2016
490 May 7;25(5):1375–84. Available from: [http://link.springer.com/10.1007/s10826-015-](http://link.springer.com/10.1007/s10826-015-0318-0)
491 0318-0
- 492 18. R. K-H, M. M, P. R, M. R. Early puberty is associated with mental health problems in
493 middle adolescence. *Soc Sci Med* [Internet]. 2003;57(6):1055–64. Available from:
494 [http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L368](http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L36876071%5Cnhttp://dx.doi.org/10.1016/S0277-9536(02)00480-X%5Cnhttp://sfxhosted.exlibrisgroup.com/galway?sid=EMBASE&issn=02779536&id=doi:10.1016%2FS0277-9536%2802%2900480-X&atitle=E)
495 [76071%5Cnhttp://dx.doi.org/10.1016/S0277-9536\(02\)00480-](http://dx.doi.org/10.1016/S0277-9536(02)00480-X%5Cnhttp://sfxhosted.exlibrisgroup.com/galway?sid=EMBASE&issn=02779536&id=doi:10.1016%2FS0277-9536%2802%2900480-X&atitle=E)
496 [X%5Cnhttp://sfxhosted.exlibrisgroup.com/galway?sid=EMBASE&issn=02779536&id](http://sfxhosted.exlibrisgroup.com/galway?sid=EMBASE&issn=02779536&id=doi:10.1016%2FS0277-9536%2802%2900480-X&atitle=E)
497 [=doi:10.1016%2FS0277-9536%2802%2900480-X&atitle=E](http://sfxhosted.exlibrisgroup.com/galway?sid=EMBASE&issn=02779536&id=doi:10.1016%2FS0277-9536%2802%2900480-X&atitle=E)
- 498 19. Dorn LD, Dahl RE, Williamson DE, Birmaher B, Axelson D, Perel J, et al.
499 Developmental Markers in Adolescence: Implications for Studies of Pubertal
500 Processes. *J Youth Adolesc*. 2003;32(5):315–24.
- 501 20. Ge X, Brody GH, Conger RD, Simons RL, Murry VMB. Contextual amplification of
502 pubertal transition effects on deviant peer affiliation and externalizing behavior among
503 African American children. *Dev Psychol*. 2002;38(1):42–54.
- 504 21. Mendle J, Beltz AM, Carter R, Dorn LD. Understanding Puberty and Its Measurement:
505 Ideas for Research in a New Generation. *J Res Adolesc*. 2019;29(1):82–95.
- 506 22. Dorn LD, Biro FM. Puberty and its measurement: A decade in review. *J Res Adolesc*.
507 2011;21(1):180–95.

“Puberty Age gap”: A new method of pubertal timing

- 508 23. Byrne ML, Whittle S, Vijayakumar N, Dennison M, Simmons JG, Allen NB. A
509 systematic review of adrenarche as a sensitive period in neurobiological development
510 and mental health. *Dev Cogn Neurosci* [Internet]. 2017 Jun;25:12–28. Available from:
511 <http://dx.doi.org/10.1016/j.dcn.2016.12.004>
- 512 24. Shirtcliff EA, Zahn-Waxler C, Klimes-Dougan B, Slattery M. Salivary
513 dehydroepiandrosterone responsiveness to social challenge in adolescents with
514 internalizing problems. *J Child Psychol Psychiatry Allied Discip*. 2007;48(6):580–91.
- 515 25. Peixoto C, José Grande A, Gomes Carrilho C, Egidio Nardi A, Cardoso A, Barciela
516 Veras A. Dehydroepiandrosterone for depressive symptoms: A systematic review and
517 meta-analysis of randomized controlled trials. *J Neurosci Res*. 2020;98(12):2510–28.
- 518 26. Han G, Miller JG, Cole PM, Zahn-Waxler C, Hastings PD. Adolescents’ internalizing
519 and externalizing problems predict their affect-specific HPA and HPG axes reactivity.
520 *Dev Psychobiol*. 2015;57(6):769–85.
- 521 27. Barendse MEA, Cosme D, Flournoy JC, Vijayakumar N, Cheng TW, Allen NB, et al.
522 Neural correlates of self-evaluation in relation to age and pubertal development in
523 early adolescent girls. *Dev Cogn Neurosci* [Internet]. 2020 Aug;44:100799. Available
524 from: <https://doi.org/10.1016/j.dcn.2020.100799>
- 525 28. Cole JH, Franke K. Predicting Age Using Neuroimaging: Innovative Brain Ageing
526 Biomarkers. *Trends Neurosci* [Internet]. 2017 Dec;40(12):681–90. Available from:
527 <http://dx.doi.org/10.1016/j.tins.2017.10.001>
- 528 29. Kaufmann T, van der Meer D, Doan NT, Schwarz E, Lund MJ, Agartz I, et al.
529 Common brain disorders are associated with heritable patterns of apparent aging of the
530 brain. *Nat Neurosci* [Internet]. 2019 Oct 24;22(10):1617–23. Available from:
531 <http://www.nature.com/articles/s41593-019-0471-7>
- 532 30. Gracia-Tabuenca Z, Moreno MB, Barrios FA, Alcauter S. Development of the brain

“Puberty Age gap”: A new method of pubertal timing

- functional connectome follows puberty-dependent nonlinear trajectories. Neuroimage
[Internet]. 2021;229(January):117769. Available from:
<https://doi.org/10.1016/j.neuroimage.2021.117769>
31. Marceau K, Ram N, Houts RM, Grimm KJ, Susman EJ. Individual differences in
boys’ and girls’ timing and tempo of puberty: Modeling development with nonlinear
growth models. Dev Psychol [Internet]. 2011 Sep;47(5):1389–409. Available from:
<http://doi.apa.org/getdoi.cfm?doi=10.1037/a0023838>
32. Rasmussen AR, Wohlfahrt-Veje C, De Renzy-Martin KT, Hagen CP, Tinggaard J,
Mouritsen A, et al. Validity of self-assessment of pubertal maturation. Pediatrics.
2015;135(1):86–93.
33. Huang B, Hillman J, Biro FM, Ding L, Dorn LD, Susman EJ. Correspondence
Between Gonadal Steroid Hormone Concentrations and Secondary Sexual
Characteristics Assessed by Clinicians, Adolescents, and Parents. J Res Adolesc.
2012;22(2):381–91.
34. Hu S, Pruessner JC, Coupé P, Collins DL. Volumetric analysis of medial temporal
lobe structures in brain development from childhood to adolescence. Neuroimage
[Internet]. 2013;74:276–87. Available from:
<http://dx.doi.org/10.1016/j.neuroimage.2013.02.032>
35. 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 (Lausanne) [Internet]. 2021 Feb 18;11(April 2020):1–22. Available from:
<https://www.frontiersin.org/articles/10.3389/fendo.2020.549928/full>
36. Gray JC, Schvey NA, Tanofsky-Kraff M. Demographic, psychological, behavioral,
and cognitive correlates of BMI in youth: Findings from the Adolescent Brain

“Puberty Age gap”: A new method of pubertal timing

- 558 Cognitive Development (ABCD) study. Psychol Med. 2020;50(9):1539–47.
- 559 37. Achenbach TM, Ruffle TM. The child behavior checklist and related forms for
560 assessing behavioral/emotional problems and competencies. Pediatr Rev.
561 2000;21(8):265–71.
- 562 38. Liang H, Zhang F, Niu X. Investigating systematic bias in brain age estimation with
563 application to post-traumatic stress disorders. Hum Brain Mapp. 2019;40(11):3143–52.
- 564 39. Rosenberg MD, Casey BJ, Holmes AJ. Prediction complements explanation in
565 understanding the developing brain. Nat Commun [Internet]. 2018;9(1):1–13.
566 Available from: <http://dx.doi.org/10.1038/s41467-018-02887-9>
- 567 40. Barendse MEA, Byrne ML, Flournoy JC, Mcneilly EA, Williamson VG, Barrett MY,
568 et al. Journal of Abnormal Psychology Internalizing Psychopathology in Early
569 Adolescent Girls Multimethod Assessment of Pubertal Timing and Associations With
570 Internalizing Psychopathology in Early Adolescent Girls. 2021;
571
572
573

“Puberty Age gap”: A new method of pubertal timing

574 Tables and Figures:

575

576 **Table 1.** Comparison of the out-of-sample prediction performance of alternative puberty age gap

577 models

	Female		Male	
	r	mae	r	mae
Combined puberty age	0.58	0.65 year	0.54	0.67 year
Hormonal puberty age	0.37	0.76 year	0.45	0.72 year
Physical development puberty age	0.55	0.67 year	0.43	0.72 year

578 r = correlation, mae = mean absolute error

579

“Puberty Age gap”: A new method of pubertal timing

580

581

582 **Figure 1.** Performance of puberty age gap models. For all three models (combined, physical

583 development and hormonal), the prediction of the age of each individual is plotted against

584 chronological age (in years). Red dots indicate individuals with a positive puberty age gap (early

585 timing) and blue dots indicate individuals with a negative puberty age gap (late timing).

586

“Puberty Age gap”: A new method of pubertal timing

587

588 **Table 2.** Comparison of the out-of-sample prediction performance of the traditional pubertal timing

589 model and the physical development puberty age model

	Female		Male	
	<i>r</i>	<i>mae</i>	<i>r</i>	<i>mae</i>
Traditional pubertal timing	0.49	0.71	0.32	0.78
Physical development Puberty age Model	0.55	0.68	0.43	0.72

590 r = correlation, mae = mean absolute error

591

“Puberty Age gap”: A new method of pubertal timing

592

593 **Table 3.** Associations between puberty age gap and mental health problems measured by linear

594 mixed effect regression

595

“Puberty Age gap”: A new method of pubertal timing

	Female						Male					
	Physical		Combined		Hormonal		Physical		Combined		Hormonal	
	t	aic	t	aic	t	aic	t	aic	t	aic	t	aic
Total	7.16***	31382.19	5.06***	31407.62	-1.61	31430.59	6.19***	34308.27	6.03***	34330.51	-0.44	34346.23
Internalising	5.81***	31145.29	3.93**	31163.47	-1.52	31176.60	6.33***	33654.96	5.29***	33677.32	-0.41	33694.64
Externalising	6.80***	30528.71	5.04***	30549.40	-1.150	30573.33	6.54***	33492.17	6.07***	33515.11	0.23	33534.59
Anxiety_Depression	1.82	26535.61	1.03	26537.87	-1.20	26537.47	5.27***	29000.46	4.78**	29019.26	-1.20	29026.61
Withdrawn_depression	6.89***	25451.83	5.67***	25467.03	0.84	25498.29	5.07***	29052.81	3.98**	29063.78	0.89	29077.61
Somatic_Complaints	4.73**	26644.49	2.68*	26659.56	-2.38	26661.07	6.35***	28592.79	5.17**	28619.53	-0.90	28632.09
Attention_problems	3.73**	26187.00	3.26*	26190.28	0.05	26200.82	2.86*	29291.46	1.89*	29296.84	-0.39	29299.45
Rule_breaking	5.38**	24226.11	3.93*	24239.47	-0.72	24254.27	6.29***	26982.27	5.31**	27005.78	-0.08	27021.57
Aggressive_Behaviour	4.82**	24915.41	3.64*	24925.27	-0.51	24938.20	5.78***	28625.15	4.46**	28646.70	-0.28	28658.31
Thought_Problems	3.10*	25920.98	2.25*	25925.51	-1.36	25928.69	5.02***	29131.42	3.61**	29147.13	-0.49	29156.23
Social_Problems	4.96**	23988.88	3.67*	24000.00	-1.24	24011.85	5.50***	26908.05	3.88**	26923.08	0.51	26937.86

“Puberty Age gap”: A new method of pubertal timing

596 Note: Statistical significance is indicated by asterisks (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; FDR corrected). The bold numbers reflect better model
597 performance.
598

“Puberty Age gap”: A new method of pubertal timing

599

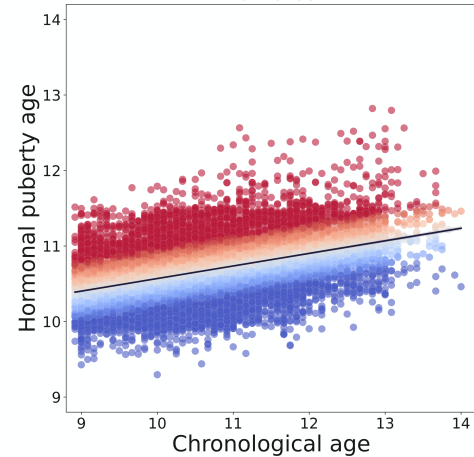
600 **Figure 2.** Bar plot showing the T-statistic from the prediction of each dimension of mental health

601 problems from different puberty age models. Hormonal puberty age gap is not significantly associated

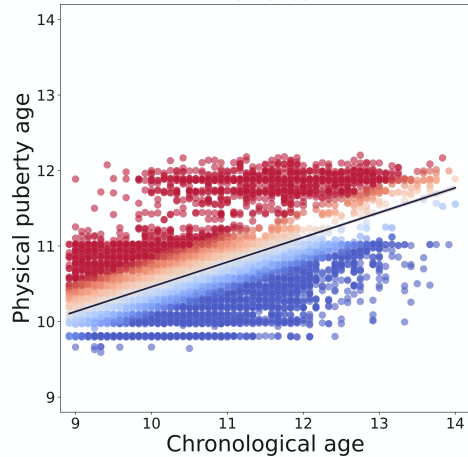
602 with any dimension of mental health problems.

603

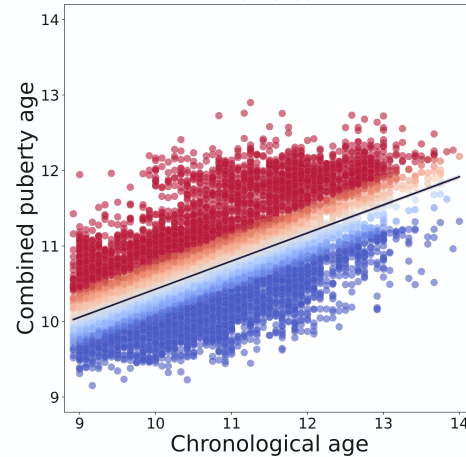
Females



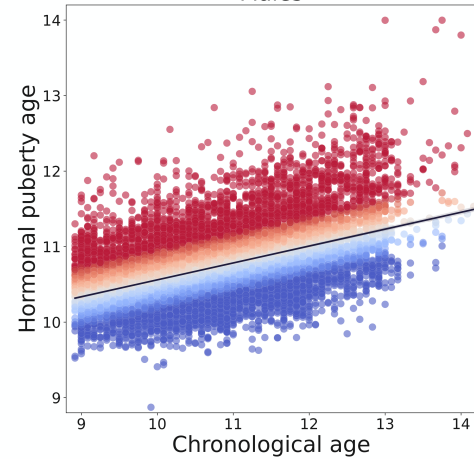
Females



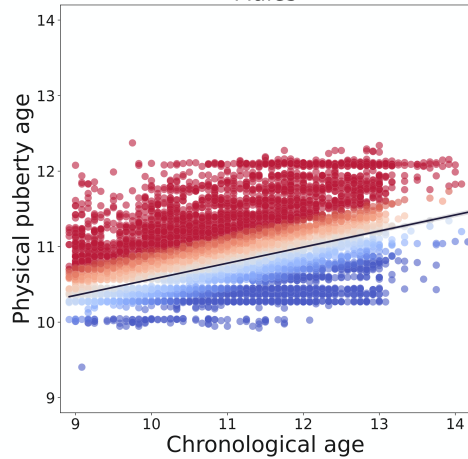
Females



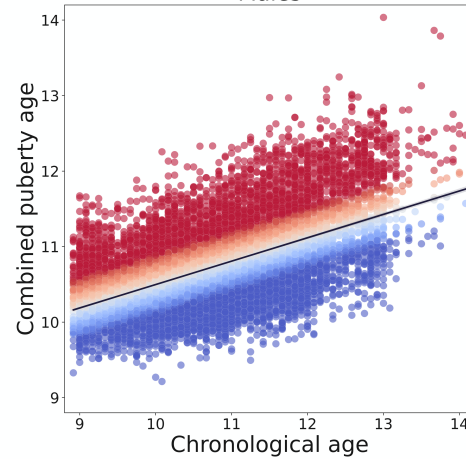
Males

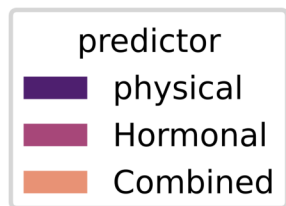


Males

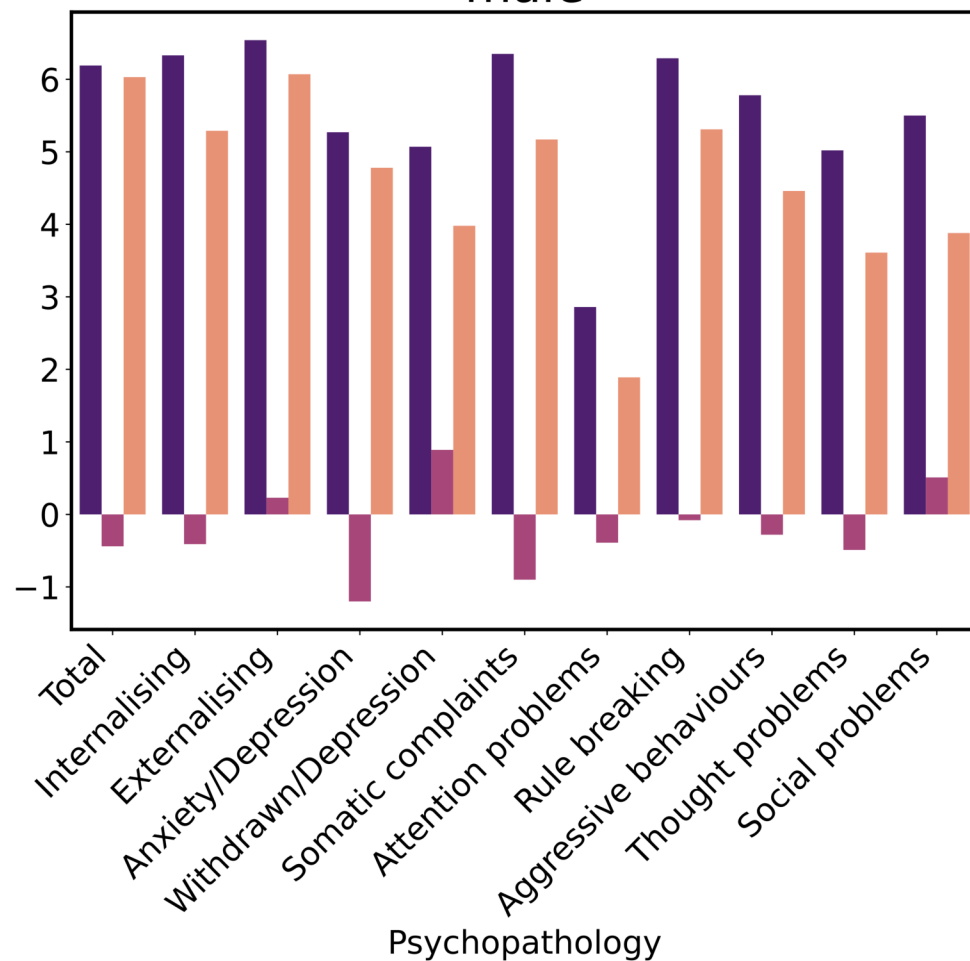


Males





male



female

