A multidimensional approach to understanding the emergence of sex differences in internalizing symptoms in adolescence

Bianca Serio [a,b,c,d,*], Robert Kohler [a], Fengdan Ye [a], Sarah D. Lichenstein [a], Sarah W. Yip [a,b]

[a] *Department of Psychiatry, Yale School of Medicine, New Haven, USA*
[b] *Child Study Center, Yale School of Medicine, New Haven, USA*
[c] *Division of Psychology and Language Sciences, University College London, London, UK*
[d] *Max Planck School of Cognition, Leipzig, Germany*

ARTICLE INFO

ABSTRACT

Women are more vulnerable to internalizing disorders (e.g., depression and anxiety). This study took an integrative developmental approach to investigate multidimensional factors associated with the emergence of sex differences in internalizing symptoms, using data from the Adolescent Brain Cognitive Development (ABCD) study. Indices of sex hormone levels (dehydroepiandrosterone, testosterone, and estradiol), physical pubertal development, task-based functional brain activity, family conflict, and internalizing symptoms were drawn from the ABCD study's baseline sample (9- to 10-year-old; $N = 11{,}844$). Principal component analysis served as a data-driven dimensionality reduction technique on the internalizing subscales to yield a single robust measure of internalizing symptoms. Moderated mediation analyses assessed whether associations between known risk factors and internalizing symptoms vary by sex. Results revealed direct and indirect effects of physical pubertal development on internalizing symptoms through family conflict across sexes. No effects were found of sex hormone levels or amygdala response to fearful faces on internalizing symptoms. Females did not report overall greater internalizing symptoms relative to males, suggesting that internalizing symptoms have not yet begun to increase in females at this age. Findings provide an essential baseline for future longitudinal research on the endocrine, neurocognitive, and psychosocial factors associated with sex differences in internalizing symptoms.

1. Introduction

Internalizing disorders—e.g., depression and anxiety—are a significant form of psychopathology that manifests during adolescence (Achenbach, 1966; Quay, 1986) and disproportionally affects women relative to men (Eaton et al., 2012; Mendle, 2014). Such sex differences in internalizing symptoms have been recorded as early as age 9 (Kessler et al., 1994), although this estimate significantly varies between studies (Breslau et al., 2017; Letcher et al., 2012; Sweeting and West, 2003; Wade et al., 2002). Studying the emergence of sex differences in internalizing symptoms from a developmental perspective and prior to peak prevalence may help to identify when sex differences in internalizing symptomatology emerge, as well as neurodevelopmental mechanisms placing females at higher risk. Given that adolescence is marked by drastic changes across multiple domains, we take a multidimensional approach to test a novel theoretical model of biological and psychosocial factors contributing to the emergence of internalizing disorders in youth, including sex hormone levels, brain function, pubertal development, and family conflict.

The emergence of differences in internalizing symptoms appears tightly linked with the onset of puberty (Soares and Zitek, 2008). Among females, more advanced pubertal development is associated with greater depression and anxiety even when controlling for age (Lewis et al., 2018; Reardon et al., 2009). Puberty is driven by steep increases in circulating sex steroid hormones, namely dehydroepiandrosterone (DHEA), testosterone, and estradiol (Mendle et al., 2019). Although there is currently no consensus on the general direction of relationships between internalizing symptoms and sex hormone levels across development, some prior studies have reported associations between increases in DHEA (Han et al., 2015; Mulligan et al., 2020), testosterone (Chronister et al., 2021; Copeland et al., 2019), and estradiol (Chronister et al., 2021; Angold et al., 1999) and increased internalizing symptoms, specifically during puberty and particularly among females.

Rising sex hormone levels during adolescence also contribute to

brain maturation, particularly the development of affective neural circuits associated with internalizing symptoms (Rubinow and Schmidt, 2019). Animal research suggests that estrogen receptors in the amygdala mediate the effects of estradiol on anxiety-like behaviors (Borrow and Handa, 2017). This preclinical work is supported by a growing body of human neuroimaging research linking the effects of DHEA (Sripada et al., 2013), testosterone (Klein et al., 2019), and estradiol (Henningsson et al., 2015) to limbic responses to emotional stimuli. Nonetheless, existing studies have focused on single-sex adult samples, leaving open questions regarding the directionality of associations in the context of emerging sex differences in puberty.

Youths show heightened amygdala reactivity to socio-emotional cues such as faces and people, as evidenced by age-related decreases in amygdala reactivity from childhood to adulthood (Bloom et al., 2022; Silvers et al., 2017). The adolescent brain is also particularly sensitive to cues from the family environment (Guyer et al., 2016). Evidence suggests that the effects of negative maternal parenting and aggression on adolescent depressive symptoms are mediated by amygdala function (Callaghan et al., 2017; Jiang et al., 2021). Therefore, alterations in amygdala activity may be a pathway through which adolescents with negative family environments develop internalizing disorders.

Pubertal development also interacts with psychosocial factors that may influence the development of internalizing symptoms. Much of the extant literature focuses on pubertal timing and suggests that adolescents with early pubertal timing are more likely to develop internalizing symptoms throughout adolescence (Hamlat et al., 2020) and adulthood (Graber et al., 2004), particularly females (Negriff and Susman, 2011). Furthermore, pubertal timing is associated with internalizing symptoms—both cross-sectionally and prospectively—when pubertal development is measured by physical assessments, but not by sex hormone levels (Barendse et al., 2020). Although this lack of significant findings for sex hormones could be partly explained by their high interpersonal variability (Mouritsen et al., 2013), associations between pubertal development and internalizing symptoms may be more meaningfully explained by psychosocial mechanisms (e.g., social reactions to the observable physical changes in one's body, such as bullying, or an individual's own reaction their changing body, such as distress), rather than biological mechanisms (i.e., the increase in hormone levels having a direct effect on increasing internalizing symptoms). For instance, the family environment strongly influences adolescent wellbeing, as underpinned by self-esteem, in different ways (Krauss et al., 2020). Family cohesion may buffer the association between internalizing behaviors and peer victimization (Jiang et al., 2016), a factor affecting self-esteem in females during puberty (Compian et al., 2009), and parental support may protect against maladjustment to peer-victimization, especially in females (Stadler et al., 2010). Conversely, harsh parenting may exacerbate the association between pubertal timing and internalizing symptoms in adolescent females (Deardorff et al., 2013). Therefore, family dynamics may interact with psychosocial stress during pubertal development, influencing internalizing symptoms and these relationships may vary between male and female adolescents.

In sum, endocrine, neurocognitive, and psychosocial factors associated with sex differences in internalizing symptoms tightly interact with one another and cannot be fully understood in isolation, particularly during adolescence, when these factors are dynamically changing. Despite observed associations between internalizing symptoms and rising sex hormone levels, affective neurodevelopment, and changing family dynamics, research is needed assessing all these risk factors concurrently in the context of emerging sex differences (Pfeifer and Allen, 2020). Furthermore, these relationships have mostly been observed in small and homogeneous samples. This study thus uses a large-scale heterogeneous dataset to investigate the putative interacting contributions of endocrine, neurocognitive, and psychosocial factors on emerging sex differences in internalizing symptoms during adolescence, as proposed in our theoretical model (Fig. 1).

This study tests our novel theoretical model in a large and diverse sample of 9- to 10-year-olds. We used heterogeneous, large-scale data from the Adolescent Brain Cognitive Development (ABCD) study, including indices of sex hormone levels, physical pubertal development, task-based functional brain activity, and family conflict. Studying 9- to 10-year-olds uniquely captures the transition from childhood to adolescence, a critical period for the emergence of sex differences in internalizing symptoms, prior to peak onset. It therefore also enables the
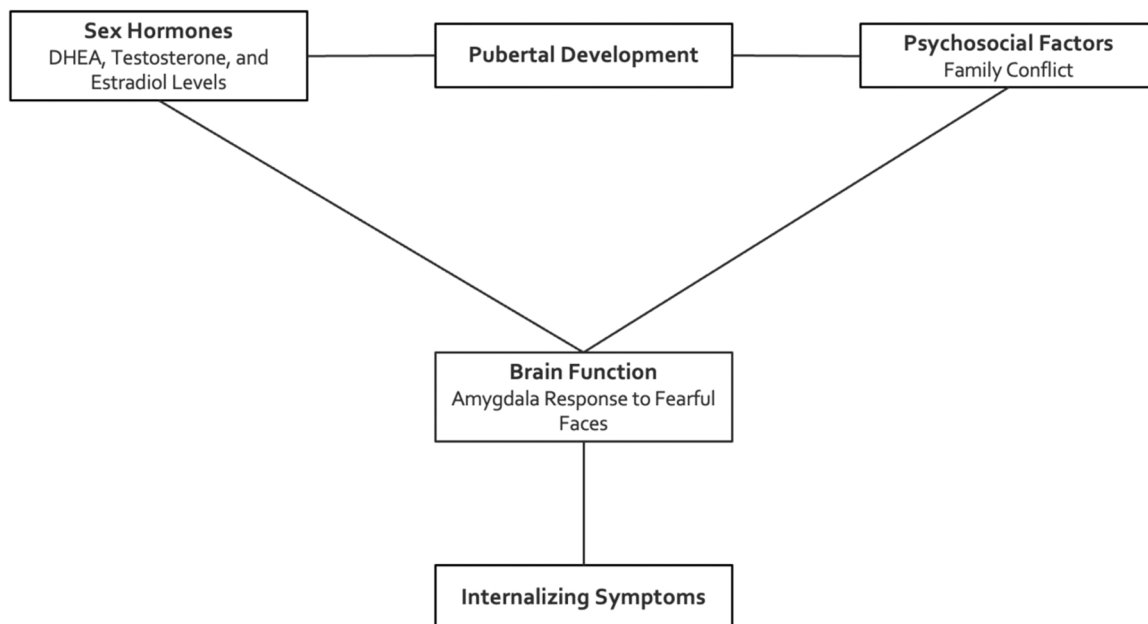

Fig. 1. Proposed theoretical model. DHEA, dehydroepiandrosterone. The proposed theoretical framework is used to test the following models: (1) effects of sex hormone levels (i.e., DHEA, testosterone, and estradiol) on internalizing symptoms through amygdala response to fearful faces; (2) effects of family conflict on internalizing symptoms through amygdala response to fearful faces; and (3) effects of physical pubertal development on internalizing symptoms through family conflict.

establishment of an empirical baseline from which relationships between endocrine, neurocognitive, and psychosocial factors may be assessed relative to internalizing symptoms in later data releases of the ABCD study. To test our proposed theoretical model, we hypothesized direct and indirect effects of (1) sex hormone levels (i.e., DHEA, testosterone, and estradiol) on internalizing symptoms through amygdala response to fearful faces; (2) family conflict on internalizing symptoms through amygdala response to fearful faces; and (3) physical pubertal development on internalizing symptoms through family conflict. We further test whether each one of these mechanisms is stronger in females relative to males in explaining sex differences in internalizing symptoms.

2. Methods

2.1. Participants

The ABCD study is the largest longitudinal consortium study of brain development in the United States (Jernigan et al., 2018), comprehensively measuring wide-ranging factors affecting adolescent neurodevelopment and health (Barch et al., 2018). Supported by the National Institute of Mental Health (NIMH), this study covers the transition from childhood to adulthood by prospectively following a baseline cohort of 9- to 10-year-olds for ten years (Volkow et al., 2018). In this study, we used the ABCD sample at baseline (Release 3.0), including 11,844 9- to 10-year-olds ($M_{age} = 118.97 \pm 7.50$ months) – see Table 1 for demographics. A detailed account of recruitment considerations and procedures is summarized in (Garavan et al., 2018; Heeringa and Berglund, 2020). Written parent consent and verbal child assent was required prior to participation (Paul et al., 2021).

2.2. Measures

2.2.1. Internalizing symptoms

Internalizing symptoms were measured with the parent/caregiver-report version of the Child Behavior Checklist (CBCL) for 6- to 18-year-old youth, a standardized questionnaire broadly assessing youth

Table 1
Demographic characteristics of the sample.

Variable	n	%
Sex		
Male	6182	52.20
Female	5662	47.80
Ethnicity		
Hispanic/Latinx	2405	20.31
Non-Hispanic/Latinx	9286	78.40
Unknown	153	1.29
Race		
White	7507	63.38
Black	1860	15.70
Asian	274	2.31
Mixed	2041	17.23
Unknown	162	1.37
Total combined household income (past 12 months)		
< $5,000	416	3.51
$5000–$11,999	419	3.54
$12,000–$15,999	274	2.31
$16,000–$24,999	522	4.41
$25000–$34,999	653	5.51
$35,000–$ 49999	932	7.87
$50,000–$74,999	1498	12.65
$75,000–$99,999	1568	13.24
$100,000–$199,999	3308	27.93
> $200,000	1238	10.45
Unknown	1016	8.58

n, number of subjects. Given that sample sizes for our different analyses varied according to data availability, the demographic characteristics reported here relate to all subjects included at least once in our analyses.

psychopathology on dimensional scales (Achenbach and Edelbrock, 1991). The CBCL includes 134 items contributing to five empirically-derived dimensions (i.e., internalizing, externalizing, thought, attention, and social symptoms), each item describing emotional and behavioral problems on 3-point Likert scales. For the current analyses, all CBCL subscales pertaining to internalizing symptoms (seven subscales) were selected in order to exhaustively capture different nuances of internalizing symptomatology, i.e., Anxious/Depressed, Withdrawn/Depressed, Somatic, DSM Depressed, DSM Anxiety, DSM Somatic, and Stress subscales. Analyses were conducted on norm-referenced t-scores (M = 50, SD = 10).

2.2.2. Sex hormone levels

The ABCD study sampled DHEA and testosterone in both males and females, and 17-b estradiol in females only. Research assistants collected salivary samples via passive drool (Uban et al., 2018). The ABCD study saliva collection protocol and sex hormone assay specifications are summarized by Herting et al. (2020). Inclusion criteria for the hormonal data were based on a decision tree adapted from Herting et al. (2020) (Fig. 2), summarizing the steps taken to conduct a quality control assessment of each replicate in order to establish a single final estimate for DHEA ($n = 10,932$) and testosterone ($n = 10,978$) levels in both sexes, and estradiol ($n = 5,119$) levels in females only.

2.2.3. Physical pubertal development

Physical pubertal development was measured with the Pubertal Development Scale (PDS), a widely used and standardized questionnaire assessing physical pubertal development (Petersen et al., 1988). It includes five items in total, assessing perceived changes in: body hair, skin, and height (both sexes); voice deepening and facial hair (males only); and breast development and menarche (females only). All items are scored on 4-point Likert scales, with higher scores indicating more advanced pubertal development, except menarche (for which 1 indicates "menstruation has not yet begun" and 4 indicates "menstruation has begun"). PDS scores were computed by averaging item scores, yielding a final score out of four. The PDS was administered to both youth and parents, but we only considered parental reports given their greater reliability and internal consistency on the PDS compared to adolescent self-reports (Carskadon and Acebo, 1993).

2.2.4. Family conflict

Family conflict was measured using the Family Environment Scale (FES), a self-report questionnaire assessing the general social climate of families (Moos and Humphrey, 1974). The questionnaire consists of nine true or false statements scored with 0 or 1. These scores were summed to yield a final score out of nine, with higher scores indicating greater family conflict.

2.2.5. fMRI emotional face N-back task

Emotional reactivity was measured with the fMRI emotional face n-back task, a validated experimental paradigm assessing emotional reactivity and working memory (Casey et al., 2018; Cohen et al., 2016), previously associated with the engagement of the fronto-amygdala circuitry (Kerestes et al., 2012). It includes both low and high memory load conditions (0-back and 2-back conditions respectively), for which subjects are required to indicate whether a given stimulus is a "match" or a "no match" to the stimulus presented at the beginning of the block (0-back condition) or two trials back (2-back condition). The task comprises two runs of eight blocks, each run consisting of four 0-back and four 2-back condition blocks. The stimuli included pictures of happy, fearful, and neutral faces, as well as places (Conley et al., 2018; Tottenham et al., 2009). Each block, for both 0-back and 2-back conditions, included 10 trials. In order to assess emotional reactivity to fearful faces, the amygdala was selected as region of interest, as measured by mean beta weights for the fearful versus neutral face contrast for both the 0-back and 2-back conditions. Data from the fMRI
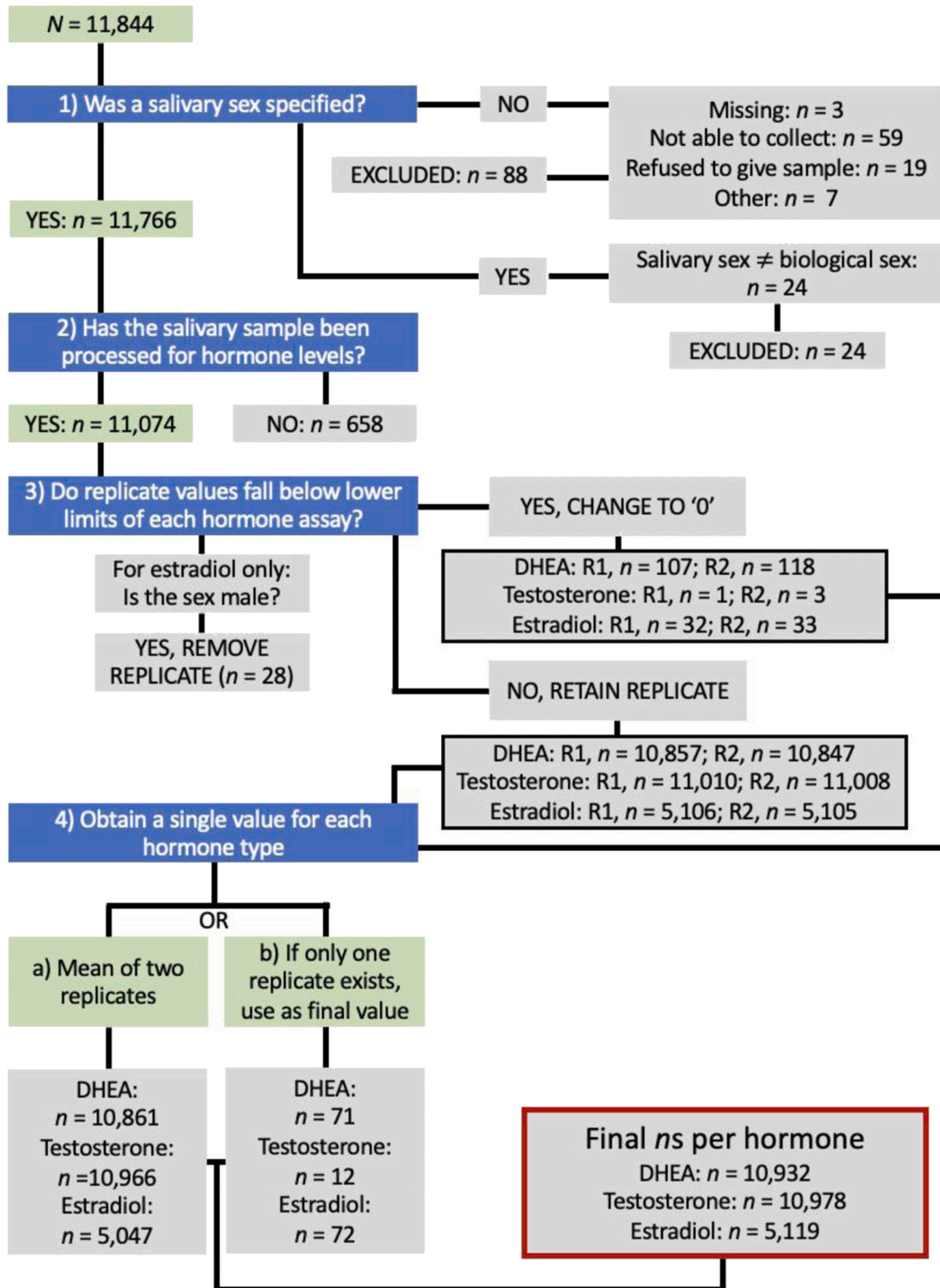
Fig. 2. Decision tree for the quality control inclusion criteria of sex hormone levels. n, number of subjects; R1, replicate 1; R2, replicate 2. Decision tree adapted from Herting et al. (2020). Data was retained if (1) salivary sex was specified and matched the biological sex; (2) the salivary sample was processed for hormone levels, i.e., at least one replicate for at least one of the three hormones was processed; (3) replicates falling below lower detection limits were valued as 0 pg/ml, and estradiol replicates measured in males were removed; (4) a final value for each hormone type was obtained by: a) calculating the mean of the two replicates; or b) if only one replicate was available, using it as the final value.

emotional face n-back task were included ($n = 7,712$) based on the recommended quality control inclusion criteria provided in ABCD Release 3.0. These criteria and the imaging procedure are outlined in the Supplementary Material.

2.3. Analytic plan

The following analyses were conducted on the ABCD study's Curated Annual Release 3.0, made publicly available on November 11th, 2020, and downloaded on February 18th, 2021, from the NIMH Data Archive (https://nda.nih.gov/abcd/query/abcd-annual-releases.html). Missing values, or data not meeting the quality control inclusion criteria for sex hormone levels and the fMRI emotional face n-back task (see Sections 2.2.2 and 2.2.5), led to the exclusion of subjects only for the affected analyses (e.g., subjects with missing sex hormone data were excluded from analyses related to sex hormones but not from other analyses). All statistical analyses were conducted in R (version 4.0.3).

2.3.1. Descriptive statistics and sex differences

Descriptive statistics were computed for all variables of interest (i.e., CBCL scores on the internalizing symptom subscales; sex hormone levels; physical pubertal development; family conflict; and amygdala response to fearful faces). Exploratory analyses assessing sex differences in all variables were conducted for descriptive purposes and to further interpret our results. Given that all data violated the normality assumption (the distributions are plotted in Supplementary Figs. S1–S12) and were mostly measured on ordinal scales, sex differences were assessed with the non-parametric Mann-Whitney U test and effect sizes were measured with rank-biserial correlation (r_{rb}) (Cureton, 1956). Bonferroni correction was used to reduce the risk of type I errors by accounting for multiple comparisons across all analyses (i.e., all variables except estradiol, measured only in females), setting the significance level to .0042 (.05/12).

2.3.2. Principal component analysis (PCA)

We used a data-driven dimensionality reduction technique (PCA) on the seven CBCL internalizing symptom subscales in order to yield a measure of internalizing psychopathology that would best reflect both the heterogeneity of internalizing symptoms and their overlap. PCA parses variance by yielding summary indices, named principal components, that are orthogonal to one another and in decreasing order of variance explained. Compared to using a total score for internalizing symptoms, treating every variable (here symptom subscales) with the same weighting, PCA yields scores based on different loadings for each variable included in the analysis, reflecting the relative weight (or "importance") of each symptom subscale. It was decided a priori that only the first principal component (PC1) would be retained as a single composite measure of internalizing symptoms to test our hypotheses.

Prior to conducting PCA, we assessed sampling adequacy of the data with the Kaiser-Meyer-Olkin (KMO) test (KMO = .76), which measures the proportion of variance among factors that may be common variance and may thus represent latent factors. Suitability for data reduction was also measured with Bartlett's test of sphericity ($\chi^2(21) = 82,637.59$, $p < .001$) by testing the null hypothesis that the correlation matrix is an identity matrix, which would imply that variables are unrelated and thus unsuitable for pattern detection. The function prcomp (stats package, R) was used to conduct PCA on the CBCL internalizing symptom subscales. All data were mean-centered and scaled.

2.3.3. Mediation and moderated mediation analyses

Moderated mediation analyses (mediation package, R) were used to test the hypothesized direct and indirect effects of: (1) sex hormone levels (i.e., DHEA and testosterone) on internalizing symptoms through amygdala response to fearful faces; (2) family conflict on internalizing symptoms through amygdala response to fearful faces; and (3) physical pubertal development on internalizing symptoms through family conflict. The direct and indirect effects of estradiol on internalizing symptoms through amygdala response to fearful faces were tested with a simple mediation analysis given that estradiol levels were only assayed in females and thus no moderation by sex could be assessed. Moderated mediation models were evaluated by first testing for an indirect effect through mediation analysis (i.e., effect of the independent variable on the dependent variable through the mediator variable), followed by testing whether sex moderates this effect (see Table 2 for a specification of the variables used to test each hypothesized model).

Conducting mediation analysis with the mediation R package requires fitting two linear regression models (Tingley et al., 2014). Model 1 specifies the effects of the independent (and moderator) variables on the mediator variable. Model 2 specifies the effects of the independent, mediator (and moderator) variables on the dependent variable. To proceed with testing for an indirect effect through mediation analysis, a statistically significant effect of the independent variable on the mediating variable (Model 1) must be observed, as well as a statistically significant effect of the mediating variable on the dependent variable when the independent and mediating variables are also included in the model as predictors (Model 2) (MacKinnon et al., 2007). Models 1 and 2 were tested for all hypotheses, including sex as the moderator variable (except when testing estradiol effects). All variables were mean-centered and scaled. Given the ABCD data's nested structure (Dick et al., 2021), we used the R function lmer (lmerTest package, R) to run generalized linear mixed-effect (GLME) models, which included random nested effects for site and family. To control for potentially confounding effects in models including hormone data, additional covariates –which have been found to influence sex hormone levels in the ABCD study– were included (i.e., collection duration (in minutes) for DHEA, testosterone, and estradiol; time of the day at sample collection (in minutes) for testosterone and estradiol; and caffeine in the past 12 h (yes/no) for estradiol) (Herting et al., 2020).

For hypotheses demonstrating statistically significant effects in Models 1 and 2, moderated mediation was tested in two steps. First, mediation was tested with the function mediate (mediation package, R) separately in males and females. Mediation was conducted with 2000 iterations of bias-corrected and accelerated (BCa) bootstrap confidence intervals (Efron, 1987) to provide more accurate estimates and correct for deviations from normality (Preacher et al., 2007). The effect size (R^2) for the mediator on the dependent variable was obtained with the function mediation (MBESS package, R). Next, moderation was tested with the function test.modmed (mediation package, R) including sex as the moderator, also with 2000 iterations of BCa bootstrap confidence intervals.

3. Results

3.1. Descriptive statistics and sex differences

Statistical results for all between-sex comparisons are shown in

Table 2
Specification of the variables used to test each hypothesized model.

IV	DV	MedV	ModV
DHEA Levels	Internalizing Symptoms	Amygdala Response to Fearful Faces	Sex
Testosterone Levels	Internalizing Symptoms	Amygdala Response to Fearful Faces	Sex
Estradiol Levels	Internalizing Symptoms	Amygdala Response to Fearful Faces	–
Family Conflict	Internalizing Symptoms	Amygdala Response to Fearful Faces	Sex
Pubertal Development	Internalizing Symptoms	Family Conflict	Sex

IV, independent variable; DV, dependent variable; MedV, mediator variable; ModV, moderator variable; DHEA, dehydroepiandrosterone.

Table 3

Mann-Whitney U test statistics for sex differences in scores on all variables of interest.

Variable	Sex		U	r_{rb}
	Males (Mean $\pm$ SD)	Females (Mean $\pm$ SD)		
CBCL Subscales (n = 11,836)				
Anxious/Depressed	53.81 $\pm$ 6.16	53.12 $\pm$ 5.71	**17,005,644**	.03
Withdrawn/Depressed	54.25 $\pm$ 6.33	52.71 $\pm$ 5.03	**15,551,276 ***	.11
Somatic	54.67 $\pm$ 5.99	55.10 $\pm$ 6.09	**18,192,098 ***	.04
DSM Depression	53.92 $\pm$ 5.97	53.25 $\pm$ 5.43	**16,371,052 ***	.06
DSM Anxiety	53.91 $\pm$ 6.23	53.05 $\pm$ 5.98	**14,094,259 ***	.19
DSM Somatic	55.32 $\pm$ 6.61	55.61 $\pm$ 6.62	17,501,550	.00
Stress	53.85 $\pm$ 6.48	52.75 $\pm$ 5.38	**15,878,058 ***	.09
Sex Hormone Levels (pg/ml)				
DHEA (n = 10,932)	54.92 $\pm$ 41.91	73.05 $\pm$ 54.96	**18,153,248 ***	.22
Testosterone (n = 10,978)	32.09 $\pm$ 19.93	36.04 $\pm$ 18.10	**17,222,647 ***	.15
Estradiol (n = 5,119)	–	1.04 $\pm$ 0.50	–	–
PDS score (n = 11,372)	1.18 $\pm$ 0.35	1.57 $\pm$ 0.56	**23,328,564 ***	.45
FES score (n = 11,817)	2.16 $\pm$ 1.97	1.93 $\pm$ 1.93	**16,125,347 ***	.07
Amygdala response to fearful faces (n = 7,712)	0.05 $\pm$ 0.49	0.06 $\pm$ 0.44	7,407,891	.00

SD, standard deviation; df, degrees of freedom; U, Mann-Whitney U test statistic; r_{rb}, rank-biserial correlation; n, number of subjects; CBCL, Child Behavior Checklist; DSM, Diagnostic and Statistical Manual of Mental Disorders; DHEA, dehydroepiandrosterone; PDS, Pubertal Development Scale; FES, Family Environment Scale. Bold denotes statistically significant sex differences at $p < .01$; * denotes statistically significant sex differences at the Bonferroni corrected significance level ($p < .0042$).

Table 3. Females did not show higher overall internalizing symptoms relative to males. Males displayed higher mean scores on the Withdrawn/Depressed ($|\mu_{Difference}| = 1.54$; $U = 15,551,276$; $p < .001$), DSM Depression ($|\mu_{Difference}| = 0.67$; $U = 16,371,052$, $p < .001$), DSM Anxiety ($|\mu_{Difference}| = 0.86$; $U = 14,094,259$, $p < .001$), and Stress ($|\mu_{Difference}| =; 1.10$; $U = 15,878,058$, $p < .001$) subscales, whereas females only showed higher mean scores on the Somatic subscale ($|\mu_{Difference}| = 0.43$, $U = 18,192,098$, $p < .001$). For all other variables of interest, except amygdala response to fearful faces, statistically significant sex differences were found. Specifically, females showed higher mean DHEA levels ($|\mu_{Difference}| = 18.13$ pg/ml; $U = 18,153,248$, $p < .001$), mean testosterone levels ($|\mu_{Difference}| = 3.95$ pg/ml, $U = 17,222,647$, $p < .001$), pubertal development ($|\mu_{Difference}| = 0.39$, $U = 23,328,564$, $p < .001$), and lower family conflict ($|\mu_{Difference}| = 0.23$, $U = 16,125,347$, $p < .001$) relative to males.

The emotional face n-back task's validity in adequately engaging the amygdala was confirmed with a one-sample t-test, showing that the beta-weight for the contrast between fearful and neutral faces in the amygdala was significantly different from 0, $t(7711) = 9.58$, $p < .001$.

3.2. Principal component analysis (PCA)

PCA yielded a first principal component (PC1) that accounted for 64.09% of the variance in all the CBCL internalizing symptom subscales. The Stress subscale contributed the most to PC1 (17.72%), followed by the Anxious/Depressed (17.24%), DSM Anxiety (16.90%), DSM

Table 4

Test statistics of the GLME models required to conduct mediation analysis for Hypothesis 3.

Model	Predictor	β	df	t	p
1	PDS	0.026	1125	2.12	**.034**
	Sex (M)	0.133	1112	6.46	**.000 ***
	PDS*Sex (M)	0.004	1128	0.20	.841
2	PDS	0.049	1122	3.75	**.000 ***
	Sex (M)	0.162	1132	7.92	**.000 ***
	FES	0.086	1022	6.59	**.000 ***
	PDS*Sex (M)	0.025	1121	1.15	.252
	FES*Sex (M)	-0.005	1013	-0.30	.765

GLME, generalized linear mixed-effect; β, standardized beta coefficient; df, degrees of freedom; PDS, Pubertal Development Scale (score); FES, Family Environment Scale (score); M, male. Bold denotes statistical significance at $p < .05$; * denotes statistical significance at $p < .001$. Model 1 predicts family conflict (FES), and Model 2 predicts internalizing symptoms (PC1 scores).

Depression (15.82%), Withdrawn/Depressed (12.45%), Somatic (11.39%), and DSM Somatic (8.46%) subscales. The variance explained by each principal component as well as the factor loadings, showing the contributions of each factor to each principal component, are summarized in Supplementary Fig. S13 and Table S14. For exploratory purposes, PCAs were also separately conducted by sex, with male and female PC1 scores showing a high correlation, $rho = 1$, $p < .001$ (Supplementary Fig. S15 and Table S16). A sensitivity analysis was also conducted to verify the effects of including overlapping CBCL subscales in our PCA (see Supplementary Material).

3.3. Mediation and moderated mediation analyses

GLME models did not indicate significant direct or indirect effects of sex hormone levels (i.e., DHEA (n = 7,138), testosterone (n = 7,157), and estradiol (n = 3,420)) on internalizing symptoms through amygdala response to fearful faces. No significant direct and indirect effects of family conflict on internalizing symptoms through amygdala response to fearful faces (n = 7,698) were found either. A statistically significant direct effect of family conflict on internalizing symptoms was observed, $t(6,758) = 5.62$, $\beta = 0.089$, $p < .001$, but this effect was not moderated by sex (as tested with the family conflict by sex interaction). Given that GLME models testing the effects of sex hormone level and family conflict on internalizing symptoms through amygdala response to fearful faces did not yield the required significance to test for a mediation effect, no additional analyses related to these hypotheses were conducted (full statistical results are available in Supplementary Table S17).

The GLME model testing direct and indirect effects of physical pubertal development on internalizing symptoms through family conflict (n = 11,346) yielded the required significant effects to test for a mediation effect, i.e., effects of pubertal development on family conflict, $t(1125) = 2.12$, $\beta = 0.026$, $p = .034$, and of family conflict on internalizing symptoms, $t(1022) = 6.59$, $\beta = 0.086$, $p < .001$, as well as a statistically significant direct effect of physical pubertal development on internalizing symptoms, $t(1122) = 3.75$, $\beta = 0.049$, $p < .001$, although this effect was not moderated by sex (as tested with the pubertal development by sex interaction; Table 4). Moderated mediation analysis revealed statistically significant indirect effects in both males and females, with an effect size for family conflict on pubertal development of $R^2 = .03\%$ (Fig. 3; Table 5). However, this indirect effect was not moderated by sex.
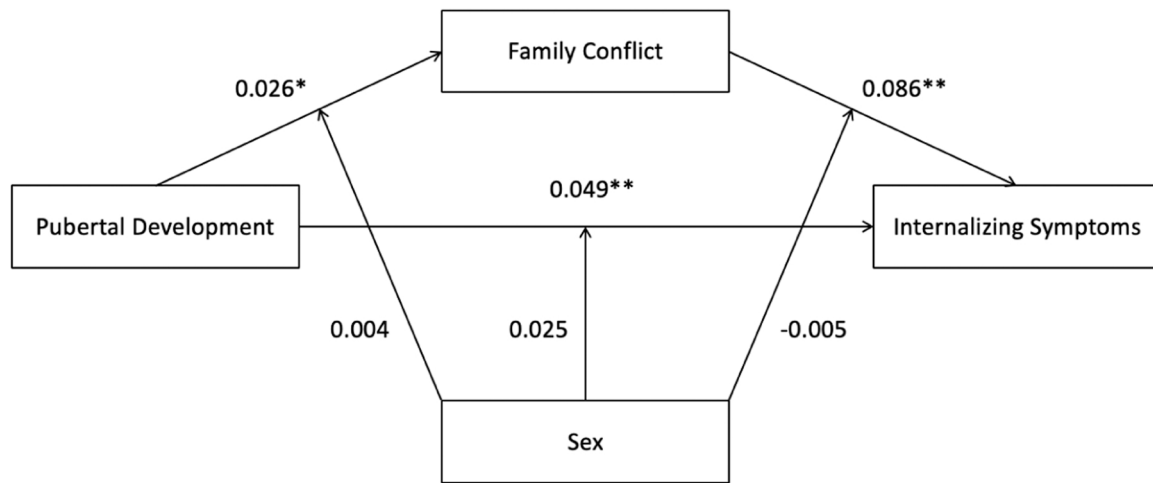
Fig. 3. Direct and indirect effects of physical pubertal development on internalizing symptoms through family conflict. Reported numbers represent standardized beta coefficients. * denotes statistically significant effects at $p < .05$, ** denotes statistically significant effects at $p < .001$.

Table 5
Test statistics of the moderated mediation analysis for Hypothesis 3.

	Mediation						Moderation		
	Males ($n = 5,928$)			Females ($n = 5,452$)					
		95% CI			95% CI			95% CI	
	β	Lower	Upper	β	Lower	Upper	β	Lower	Upper
PDS → FES → internalizing symptoms (indirect effect)	**0.005**	0.002	0.010	**0.004***	0.002	0.010	0.000	-0.004	0.005

β, standardized beta coefficient; PDS, Pubertal Development Scale (score); FES, Family Environment Scale (score); CI, confidence interval. Bold denotes statistical significance at $p < .05$; * denotes statistical significance at $p < .001$.

4. Discussion

This is the largest study to date taking a developmental approach to investigate the mechanisms involved in the emergence of internalizing psychopathology, including endocrine, neurocognitive, and psychosocial factors. We aimed to capture the emergence of sex differences in internalizing symptoms in order to elucidate the factors placing females at greater risk of internalizing psychopathology. We used heterogeneous data including indices of sex hormone levels, physical pubertal development, task-based functional brain activity, and family conflict to test a novel theoretical model (Fig. 1) in the ABCD study's large and diverse 9- to 10-year-old sample. Mediation analyses indicated significant direct and indirect effects of physical pubertal development on internalizing symptoms through family conflict, although these effects were not moderated by sex. In fact, females did not report overall higher levels of internalizing symptoms. Furthermore, no endocrine or neurocognitive effects were observed. Given that adolescence is a dynamic developmental period, our findings broadly encourage testing our model longitudinally, specifically at later stages of puberty. This first analysis provides an essential empirical baseline from which relationships between endocrine, neurocognitive, and psychosocial factors may be assessed relative to internalizing symptoms in later data releases of the ABCD study.

4.1. Greater prevalence of internalizing symptoms in females is not yet observable at ages 9–10

Females in the ABCD sample did not report overall greater internalizing symptoms, scoring higher than males on only one out of seven CBCL subscales (i.e., somatic symptoms – consistent with previous findings (Ruchkin and Schwab-Stone, 2014)). While sex differences in CBCL subscales were modest at this early timepoint, these results

nonetheless suggest that future research examining internalizing symptom subtypes may help to explain inconsistencies in previous literature on the presence and directionality of sex differences in pre-adolescents (Kessler et al., 1994; Breslau et al., 2017; Letcher et al., 2012; Sweeting and West, 2003; Wade et al., 2002). Prior studies have typically relied on overall scores for internalizing symptoms rather than systematically examining symptom subtypes, which may be a potential reason for mixed results. Furthermore, the skewed distribution of PDS scores suggests that the sample has largely not yet entered puberty or is only at its beginning. Therefore, in addition to suggesting that greater prevalence of internalizing symptoms in females may not yet be detectable at ages 9–10 and at such an early stage of pubertal development, our findings on the differing directionality of sex differences according to symptom subtype also suggest the importance of considering patterns in symptomatology when studying sex differences in internalizing psychopathology.

4.2. Pubertal development is associated with internalizing symptoms through family conflict across sexes

As hypothesized in our theoretical model, physical pubertal development was positively associated with internalizing symptoms both directly and indirectly—through family conflict—across sexes, implying that family conflict may amplify pubertal effects on internalizing symptoms. These findings align with previous research suggesting that harsh parenting exacerbates pubertal timing effects on internalizing symptoms in females (Deardorff et al., 2013). Our results thus extend previous findings by establishing indirect effects in both sexes, as early as ages 9–10, generalizing effects to pubertal development rather than pubertal timing, and more generally regarding family conflict rather than harsh parenting specifically.

Contrary to our theoretical model, associations between pubertal

development, family conflict, and internalizing symptoms were not moderated by sex. At ages 9–10 in the ABCD sample, family conflict accounted for internalizing symptoms similarly in males and females, consistent with evidence suggesting that sex differences in family effects on internalizing symptoms do not emerge until mid-adolescence (Crawford et al., 2001). Furthermore, FES scores in our sample were lower in females, suggesting less family conflict. This aligns with females typically reporting greater family cohesion than males (Sze et al., 2013), which may heighten their long-term sensitivity to conflict and trigger internalizing symptoms at later pubertal stages (Lewis et al., 2015), as family conflict increases throughout adolescence (Steinberg, 1988). Therefore, the strength of the association between pubertal development and internalizing symptoms through family conflict is expected to increase over time, especially in females, and may contribute to the emergence of sex differences in these effects.

4.3. Endocrine and neurocognitive effects on internalizing symptoms are not yet observable at ages 9–10

Further contrary to our hypothesized theoretical model, our findings do not indicate effects of sex hormone levels and family conflict on internalizing symptoms through amygdala response to fearful faces. The ABCD sample at ages 9–10 may thus be too young to show the mechanisms proposed in our theoretical model, including associated sex differences, as the trend of greater internalizing symptoms in females is not yet observable in this sample. In fact, no sex differences in amygdala response to fearful faces were detected either, an effect consistently reported in adults (Andreano et al., 2014). Furthermore, no associations were found between amygdala response to fearful faces and internalizing symptoms, common to all unsupported hypotheses in our study attempting to establish brain function as the pathway from risk factors to internalizing symptoms (Drevets et al., 2008). Research suggests that heightened amygdala response to fearful faces in adult females, but not males, may contribute to females' greater vulnerability to internalizing disorders (Dickie and Armony, 2008). This association has however not yet been established in youth, consistent with our null findings. Previous research on adolescent depression also found no sex differences in amygdala response to negative stimuli at ages 13–17 (Yang et al., 2010). Therefore, sex differences in amygdala reactivity to negative stimuli may only emerge in late-adolescence/early-adulthood as a potential risk factor for depression in females, and our sample may be too young to capture this effect.

It may also be too early, at ages 9–10, to observe direct effects of hormone levels on internalizing symptoms. Our findings did indicate higher levels of DHEA and testosterone in females relative to males, consistent with previous research suggesting that females begin puberty before males (Marshall, 1978). However, sex differences in hormone levels steadily increase throughout adolescence and endocrine effects on internalizing symptoms may thus take some time to emerge. Menarche has also been associated with the development of internalizing symptoms (Mendle et al., 2018), particularly in females with heightened neurological vulnerability to fluctuating ovarian hormone levels (Barth et al., 2015). However, these effects could not be reliably assessed in our sample given the small proportion of females reporting menarche. Nevertheless, as this was the first study exploring endocrine effects at such a young age, our null findings are still valuable in suggesting that it may be too early, at ages 9–10, to observe endocrine effects on internalizing symptoms.

4.4. Methodological considerations

Although we found some statistically significant results, effect sizes were small by traditional standards, e.g., $R^2 = .03\%$ for the effect of family conflict on physical pubertal development. Such small effects may have multiple explanations. Firstly, research conducted on large samples (Miller et al., 2016), including the ABCD sample (Marek et al.,

2020), has recently shown similarly small effect sizes, e.g., statistically significant correlations between parental acceptance and total psychiatric problems ($r = -.09$) or between caffeine consumption and sleep problems ($r = .04$) (Owens et al., 2021). Established benchmarks for interpreting effect sizes may thus be historically based on underpowered or biased studies overestimating effect sizes, and should therefore not be used to assess the meaningfulness of our effects (Schäfer and Schwarz, 2019). Secondly, large-scale research initiatives are powered to detect small effects, which are suggested to more closely approximate true population values (Dick et al., 2021). Small effect sizes in large, multisite studies of heterogeneous samples may also be due to increased statistical noise diluting effect sizes (Feaster et al., 2011). Nevertheless, large and diverse samples, such as the ABCD cohort, directly address the issue of low power and bias in research (Ioannidis, 2005), and are thus important to overcome the typically low reproducibility of findings in psychology and neuroscience (Button et al., 2013). Thirdly, given that our sample is not clinical, we are likely looking at subtle variation in internalizing pathology within the normal range and would not expect to observe large effect sizes in this context and at this age. However, in developmental psychopathology, effects are expected to accumulate and grow over time (Wenar and Kerig, 2000). Small effects like ours may thus additively explain a substantial proportion of variance in neurodevelopmental trajectories (Boyle et al., 2017).

4.4.1. Limitations

Although this study has valuable strengths, some methodological and conceptual limitations should be acknowledged. Firstly, the fMRI emotional face n-back task may not adequately capture emotional reactivity due to potential interference from the tasks' working memory component. Motivated to perform well and focusing on evaluating whether stimuli matched those presented two trials back, subjects may have not paid explicit attention to the emotional valence of the presented stimuli. In fact, the construction of conscious emotional experience is argued to require the integration of both implicit (e.g., visceral) and explicit (e.g., attentional) emotional processes (Quirin and Lane, 2012). The emotional face n-back task may thus not be the most suitable measure of emotional reactivity. Nevertheless, our data shows evidence that the task did yield differential amygdala activations in the contrast between fearful and neutral faces, supporting its intended use to establish the neurocognitive effects depicted by our theoretical model.

Secondly, as our analyses were conducted on data collected at a single timepoint, no inferences on the directionality of direct and indirect effects could be made. Mediation analysis indeed requires longitudinal data to establish directionality (Maxwell and Cole, 2007). However, mediation models tested with single timepoint data have shown acceptable performance in a recent study using simulations to compare different mediation models including an equal amount of data but varying numbers of timepoints (Cain et al., 2018). Therefore, mediation analysis still serves our aims of gaining a preliminary mechanistic understanding of how multilevel data may interact to yield internalizing symptoms through direct and indirect effects.

Thirdly, the literature heavily focuses on effects of pubertal timing and tempo on internalizing psychopathology (Ullsperger and Nikolas, 2017). However, this specific phenomenon is not investigated in the current study for multiple reasons. Firstly, given the novelty and complexity of our hypothesized effects, we focused on the basic phenomenon that entering puberty per se increases the risk of internalizing disorders. This effect, and its interactions with the other heterogeneous risk factors under study, should be understood before investigating the additionally complex role that pubertal timing and tempo may play in this framework. Secondly, our sample is too young to adequately study pubertal timing and tempo. Many subjects' PDS scores indicated that they have not yet entered puberty, making it virtually impossible to test the hypothesis of accelerated pubertal development. Finally, pubertal timing and tempo are best investigated in a longitudinal design with multiple timepoints to track the individual trajectories of pubertal

development, further capturing individual differences in pubertal onset (as suggested in Section 4.4).

4.5. Implications and future directions

This is the first study taking a developmental perspective to concurrently investigate endocrine, neurocognitive, and psychosocial factors associated with emerging sex differences in internalizing symptoms in a 9- to 10-year-old sample. By taking an integrative approach, we considered the additive contributions of heterogeneous risk factors on internalizing psychopathology and how they may interact with one another.

Our findings have theoretical implications in suggesting that sex differences in internalizing symptoms are not yet present among 9- to 10-year-olds. Considering the ABCD sample's size and diversity, these findings valuably contribute to existing mixed research on the age at which sex differences emerge. Furthermore, null findings for direct and indirect effects of sex hormones and family conflict on internalizing symptoms through brain function suggest that it may also be too early, at ages 9–10, to observe interactive effects of endocrine, neurocognitive, and psychosocial risk factors.

In the field of developmental psychopathology, null findings serve as an important baseline for future longitudinal research (Luna et al., 2021). Given that adolescence is a transitional developmental period, the presently hypothesized effects and associated sex differences are expected to emerge and grow over time (Casey et al., 2008). As the ABCD study aims to prospectively follow its baseline cohort for ten years, it will provide an unprecedented opportunity to build on the present findings by testing our proposed theoretical model at different ages. Longitudinal investigation of the developmental trajectories of sex differences in internalizing symptoms will further allow us to assess the directionality of effects.

Future research on emerging sex differences in internalizing symptoms should also investigate fronto-limbic network connectivity rather than amygdala reactivity alone, given recent findings identifying sex differences in associations between amygdala functional connectivity and internalizing symptoms in neurotypical youth (Padgaonkar et al., 2020). In addition, psychosocial effects of peer relationships need to be considered in addition to family relationships, as youth dynamically shift from one to the other over the course of adolescence (Larson et al., 1996). Furthermore, the effects of pubertal timing and tempo on risk for internalizing psychopathology (Ullsperger and Nikolas, 2017), as well as the psychological impact of individual differences in pubertal timing and how these effects may interact with endocrine and neurocognitive risk factors, should be studied over time. The longitudinal investigation of our theoretical model with the proposed additional considerations may have major clinical implications for the early diagnosis and detection of youth at risk by informing more targeted interventions for prevention and treatment.

5. Conclusion

Taking a developmental and multidimensional approach is necessary to understand the greater prevalence of internalizing symptoms in females. We concurrently assessed contributions from endocrine, neurocognitive, and psychosocial perspectives to study interacting mechanisms underlying emerging sex differences in a large and diverse sample of 9- to 10-year-olds. Consistent with previous research, effects of pubertal development on internalizing symptoms through family conflict were observed across sexes. However, it seems too early at this age to observe the surge of internalizing symptoms in females, as well as interactive effects of endocrine, neurocognitive, and psychosocial factors. By examining mechanisms of risk prior to the peak period for the onset of psychopathology, our findings provide an important baseline for future longitudinal research on the emergence of sex differences in internalizing symptoms.

Appendix A. Supplementary material

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.dcn.2022.101182.

References

Achenbach, T.M., 1966. The classification of children's psychiatric symptoms: a factor-analytic study. Psychol. Monogr. Gen. Appl. 80 (7), 1.

Achenbach, T.M., Edelbrock, C., 1991. Child Behavior Checklist. Burlington (Vt) 7, pp. 371–392.

Andreano, J.M., Dickerson, B.C., Barrett, L.F., 2014. Sex differences in the persistence of the amygdala response to negative material. Soc. Cogn. Affect. Neurosci. 9 (9), 1388–1394.

Angold, A., et al., 1999. Pubertal changes in hormone levels and depression in girls. Psychol. Med. 29 (5), 1043–1053.

Barch, D.M., et al., 2018. Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: rationale and description. Dev. Cogn. Neurosci. 32, 55–66.

Barendse, M., et al., 2020. Multi-Method Assessment of Pubertal Timing and Associations with Internalizing Psychopathology in Early Adolescent Girls. PsyArXiv Preprints.

Barth, C., Villringer, A., Sacher, J., 2015. Sex hormones affect neurotransmitters and shape the adult female brain during hormonal transition periods. Front. Neurosci. 9, 37.

Bloom, P.A., et al., 2022. Age-related change in task-evoked amygdala—prefrontal circuitry: a multiverse approach with an accelerated longitudinal cohort aged 4–22 years. Hum. Brain Mapp.

Borrow, A., Handa, R.J., 2017. Estrogen receptors modulation of anxiety-like behavior. Vitam. Horm. 103, 27–52.

Boyle, E.A., Li, Y.I., Pritchard, J.K., 2017. An expanded view of complex traits: from polygenic to omnigenic. Cell 169 (7), 1177–1186.

Breslau, J., et al., 2017. Sex differences in recent first-onset depression in an epidemiological sample of adolescents. Transl. Psychiatry 7 (5) e1139-e1139.

Button, K.S., et al., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14 (5), 365–376.

Cain, M.K., Zhang, Z., Bergeman, C., 2018. Time and other considerations in mediation design. Educ. Psychol. Meas. 78 (6), 952–972.

Callaghan, B.L., et al., 2017. Amygdala resting connectivity mediates association between maternal aggression and adolescent major depression: a 7-year longitudinal study. J. Am. Acad. Child Adolesc. Psychiatry 56 (11), 983–991.

Carskadon, M.A., Acebo, C., 1993. A self-administered rating scale for pubertal development. J. Adolesc. Health 14 (3), 190–195.

Casey, B., et al., 2018. The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. Dev. Cogn. Neurosci. 32, 43–54.

Casey, B., Getz, S., Galvan, A., 2008. The adolescent brain. Dev. Rev. 28 (1), 62–77.

Chronister, B.N., et al., 2021. The effects of testosterone, estradiol, dehydroepiandrosterone, and cortisol on anxiety and depression symptoms in ecuadorian adolescents. Am. Acad. Pediatr. 147, 3.

Cohen, A., et al., 2016. The Impact of Emotional Cues on Short-term and Long-term Memory during Adolescence. Program No. 90.25 Neuroscience Meeting Planner. Society for Neuroscience, San Diego, CA.

Compian, L.J., Gowen, L.K., Hayward, C., 2009. The interactive effects of puberty and peer victimization on weight concerns and depression symptoms among early adolescent girls. J. Early Adolesc. 29 (3), 357–375.

Conley, M.I., et al., 2018. The racially diverse affective expression (RADIATE) face stimulus set. Psychiatry Res. 270, 1059–1067.

Copeland, W.E., et al., 2019. Early pubertal timing and testosterone associated with higher levels of adolescent depression in girls. J. Am. Acad. Child Adolesc. Psychiatry 58 (12), 1197–1206.

Crawford, T.N., et al., 2001. Internalizing symptoms in adolescents: gender differences in vulnerability to parental distress and discord. J. Res. Adolesc. 11 (1), 95–118.

Cureton, E.E., 1956. Rank-biserial correlation. Psychometrika 21 (3), 287–290.

Deardorff, J., et al., 2013. Pubertal timing and Mexican-origin girls' internalizing and externalizing symptoms: the influence of harsh parenting. Dev. Psychol. 49 (9), 1790.

Dick, A.S., et al., 2021. Meaningful associations in the adolescent brain cognitive development study. NeuroImage, 118262.

Dickie, E.W., Armony, J.L., 2008. Amygdala responses to unattended fearful faces: interaction between sex and trait anxiety. Psychiatry Res. Neuroimaging 162 (1), 51–57.

Drevets, W.C., Price, J.L., Furey, M.L., 2008. Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. Brain Struct. Funct. 213 (1–2), 93–118.

Eaton, N.R., et al., 2012. An invariant dimensional liability model of gender differences in mental disorder prevalence: evidence from a national sample. J. Abnorm. Psychol. 121 (1), 282.

Efron, B., 1987. Better bootstrap confidence intervals. J. Am. Stat. Assoc. 82 (397), 171–185.

Feaster, D.J., Mikulich-Gilbertson, S., Brincks, A.M., 2011. Modeling site effects in the design and analysis of multi-site trials. Am. J. Drug Alcohol Abus. 37 (5), 383–391.

Garavan, H., et al., 2018. Recruiting the ABCD sample: design considerations and procedures. Dev. Cogn. Neurosci. 32, 16–22.

Graber, J.A., et al., 2004. Is pubertal timing associated with psychopathology in young adulthood? J. Am. Acad. Child Adolesc. Psychiatry 43 (6), 718–726.

Guyer, A.E., Silk, J.S., Nelson, E.E., 2016. The neurobiology of the emotional adolescent: from the inside out. Neurosci. Biobehav. Rev. 70, 74–85.

Hamlat, E.J., et al., 2020. Early pubertal timing predicts onset and recurrence of depressive episodes in boys and girls. J. Child Psychol. Psychiatry 61 (11), 1266–1274.

Han, G., et al., 2015. Adolescents' internalizing and externalizing problems predict their affect-specific HPA and HPG axes reactivity. Dev. Psychobiol. 57 (6), 769–785.

Heeringa, S.G., Berglund, P.A., 2020. A Guide for Population-based Analysis of the Adolescent Brain Cognitive Development (ABCD) Study Baseline Data. BioRxiv.

Henningsson, S., et al., 2015. Role of emotional processing in depressive responses to sex-hormone manipulation: a pharmacological fMRI study. Transl. Psychiatry 5 (12) e688-e688.

Herting, M.M., et al., 2020. Correspondence between perceived pubertal development and hormone levels in 9-10 year-olds from the Adolescent Brain Cognitive Development study. Front. Endocrinol. 11.

Ioannidis, J.P., 2005. Why most published research findings are false. PLoS Med. 2 (8), e124.

Jernigan, T.L., Brown, S.A., Dowling, G.J., 2018. The adolescent brain cognitive development study. J. Res. Adolesc.: Off. J. Soc. Res. Adolesc. 28 (1), 154.

Jiang, N., et al., 2021. Negative parenting affects adolescent internalizing symptoms through alterations in amygdala-prefrontal circuitry: a longitudinal twin study. Biol. Psychiatry 89 (6), 560–569.

Jiang, Y., et al., 2016. Buffering the effects of peer victimization on adolescent non-suicidal self-injury: the role of self-compassion and family cohesion. J. Adolesc. 53, 107–115.

Kerestes, R., et al., 2012. Abnormal prefrontal activity subserving attentional control of emotion in remitted depressed patients during a working memory task with emotional distracters. Psychol. Med. 42 (1), 29.

Kessler, R.C., et al., 1994. Sex and depression in the National Comorbidity Survey. II: cohort effects. J. Affect. Disord. 30 (1), 15–26.

Klein, S., et al., 2019. Increased neural reactivity to emotional pictures in men with high hair testosterone concentrations. Soc. Cogn. Affect. Neurosci. 14 (9), 1009–1016.

Krauss, S., Orth, U., Robins, R.W., 2020. Family environment and self-esteem development: a longitudinal study from age 10 to 16. J. Personal. Soc. Psychol. 119 (2), 457.

Larson, R.W., et al., 1996. Changes in adolescents' daily interactions with their families from ages 10 to 18: disengagement and transformation. Dev. Psychol. 32 (4), 744.

Letcher, P., et al., 2012. Precursors and correlates of anxiety trajectories from late childhood to late adolescence. J. Clin. Child Adolesc. Psychol. 41 (4), 417–432.

Lewis, A., et al., 2015. Gender differences in adolescent depression: differential female susceptibility to stressors affecting family functioning. Aust. J. Psychol. 67 (3), 131–139.

Lewis, G., et al., 2018. The association between pubertal status and depressive symptoms and diagnoses in adolescent females: a population-based cohort study. PloS One 13 (6), e0198804.

Luna, B., Tervo-Clemmens, B., Calabro, F.J., 2021. Considerations when characterizing adolescent neurocognitive development. Biol. Psychiatry 89 (2), 96–98.

MacKinnon, D.P., Fairchild, A.J., Fritz, M.S., 2007. Mediation analysis. Annu. Rev. Psychol. 58, 593–614.

Marek, S., et al., 2020. Towards Reproducible Brain-wide Association Studies. BioRxiv.

Marshall, W.A., 1978. Puberty. In: Falkner, F., Tanner, J.M. (Eds.), Human Growth. Springer, New York, pp. 141–181.

Maxwell, S.E., Cole, D.A., 2007. Bias in cross-sectional analyses of longitudinal mediation. Psychol. Methods 12 (1), 23.

Mendle, J., 2014. Why puberty matters for psychopathology. Child Dev. Perspect. 8 (4), 218–222.

Mendle, J., et al., 2019. Understanding puberty and its measurement: ideas for research in a new generation. J. Res. Adolesc. 29 (1), 82–95.

Mendle, J., Ryan, R.M., McKone, K.M., 2018. Age at menarche, depression, and antisocial behavior in adulthood. Pediatrics 141, 1.

Miller, K.L., et al., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19 (11), 1523–1536.

Moos, R.H., Humphrey, B., 1974. Preliminary Manual for Family Environment Scale, Work Environment Scale, Group Environment Scale. Consulting Psychologists Press, Palo Alto, CA.

Mouritsen, A., et al., 2013. The pubertal transition in 179 healthy Danish children: associations between pubarche, adrenarche, gonadarche, and body composition. Eur. J. Endocrinol. 168 (2), 129–136.

Mulligan, E.M., et al., 2020. Increased dehydroepiandrosterone (DHEA) is associated with anxiety in adolescent girls. Psychoneuroendocrinology 119, 104751.

Negriff, S., Susman, E.J., 2011. Pubertal timing, depression, and externalizing problems: a framework, review, and examination of gender differences. J. Res. Adolesc. 21 (3), 717–746.

Owens, M.M., et al., 2021. Recalibrating expectations about effect size: a multi-method survey of effect sizes in the ABCD study. PLoS ONE 16 (9), e0257535.

Padgaonkar, N., et al., 2020. Sex differences in internalizing symptoms and amygdala functional connectivity in neurotypical youth. Dev. Cogn. Neurosci. 44, 100797.

Paul, S.E., et al., 2021. Associations between prenatal cannabis exposure and childhood outcomes: results from the ABCD study. JAMA Psychiatry 78 (1), 64–76.

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

Pfeifer, J.H., Allen, N.B., 2020. Puberty initiates cascading relationships between neurodevelopmental, social, and internalizing processes across adolescence. Biol. Psychiatry 89 (2), 99–108.

Preacher, K.J., Rucker, D.D., Hayes, A.F., 2007. Addressing moderated mediation hypotheses: theory, methods, and prescriptions. Multivar. Behav. Res. 42 (1), 185–227.

Quay, H.C., 1986. Classification. In: Quay, H.C., Werry, J.S. (Eds.), Psychopathological Disorders of Childhood. Wiley, New York, pp. 1–34.

Quirin, M., Lane, R.D., 2012. The construction of emotional experience requires the integration of implicit and explicit emotional processes. Behav. Brain Sci. 35 (3), 159.

Reardon, L.E., Leen-Feldner, E.W., Hayward, C., 2009. A critical review of the empirical literature on the relation between anxiety and puberty. Clin. Psychol. Rev. 29 (1), 1–23.

Rubinow, D.R., Schmidt, P.J., 2019. Sex differences and the neurobiology of affective disorders. Neuropsychopharmacology 44 (1), 111–128.

Ruchkin, V., Schwab-Stone, M., 2014. A longitudinal study of somatic complaints in urban adolescents: the role of internalizing psychopathology and somatic anxiety. J. Youth Adolesc. 43 (5), 834–845.

Schäfer, T., Schwarz, M.A., 2019. The meaningfulness of effect sizes in psychological research: differences between sub-disciplines and the impact of potential biases. Front. Psychol. 10, 813.

Silvers, J.A., et al., 2017. The transition from childhood to adolescence is marked by a general decrease in amygdala reactivity and an affect-specific ventral-to-dorsal shift in medial prefrontal recruitment. Dev. Cogn. Neurosci. 25, 128–137.

Soares, C.N., Zitek, B., 2008. Reproductive hormone sensitivity and risk for depression across the female life cycle: a continuum of vulnerability? J. Psychiatry Neurosci. 33 (4), 331–343.

Sripada, R.K., et al., 2013. DHEA enhances emotion regulation neurocircuits and modulates memory for emotional stimuli. Neuropsychopharmacology 38 (9), 1798–1807.

Stadler, C., et al., 2010. Peer-victimization and mental health problems in adolescents: are parental and school support protective? Child Psychiatry Hum. Dev. 41 (4), 371–386.

Steinberg, L., 1988. Reciprocal relation between parent-child distance and pubertal maturation. Dev. Psychol. 24 (1), 122.

Sweeting, H., West, P., 2003. Sex differences in health at ages 11, 13 and 15. Soc. Sci. Med. 56 (1), 31–39.

Sze, T.-M., et al., 2013. Sex differences in the development of perceived family cohesion and depressive symptoms in Taiwanese adolescents. Psychol. Rep. 113 (1), 54–72.

Tingley, D., et al., 2014. Mediation: R package for causal mediation analysis. J. Stat. Softw. 59, 5.

Tottenham, N., et al., 2009. The NimStim set of facial expressions: judgments from untrained research participants. Psychiatry Res. 168 (3), 242–249.

Uban, K.A., et al., 2018. Biospecimens and the ABCD study: rationale, methods of collection, measurement and early data. Dev. Cogn. Neurosci. 32, 97–106.

Ullsperger, J.M., Nikolas, M.A., 2017. A meta-analytic review of the association between pubertal timing and psychopathology in adolescence: are there sex differences in risk? Psychol. Bull. 143 (9), 903.

Volkow, N.D., et al., 2018. The conception of the ABCD study: from substance use to a broad NIH collaboration. Dev. Cogn. Neurosci. 32, 4–7.

Wade, T.J., Cairney, J., Pevalin, D.J., 2002. Emergence of gender differences in depression during adolescence: national panel results from three countries. J. Am. Acad. Child Adolesc. Psychiatry 41 (2), 190–198.

Wenar, C., Kerig, P., 2000. Developmental Psychopathology: From Infancy through Adolescence, fourth ed. McGraw-Hill.

Yang, T.T., et al., 2010. Adolescents with major depression demonstrate increased amygdala activation. J. Am. Acad. Child Adolesc. Psychiatry 49 (1), 42–51.