NEW RESEARCH

Check for updates

Irritability Trajectories, Cortical Thickness, and Clinical Outcomes in a Sample Enriched for Preschool Depression

David Pagliaccio, PhD, Daniel S. Pine, MD, Deanna M. Barch, PhD, Joan L. Luby, MD, Ellen Leibenluft, MD

Objective: Cross-sectional, longitudinal, and genetic associations exist between irritability and depression. Prior studies have examined developmental trajectories of irritability, clinical outcomes, and associations with child and familial depression. However, studies have not integrated neurobiological measures. The present study examined developmental trajectories of irritability, clinical outcomes, and cortical structure among preschoolers oversampled for depressive symptoms.

Method: Beginning at 3 to 5 years old, a sample of 271 children enriched for early depressive symptoms were assessed longitudinally by clinical interview. Latent class mixture models identified trajectories of irritability severity. Risk factors, clinical outcomes, and cortical thickness were compared across trajectory classes. Cortical thickness measures were extracted from 3 waves of magnetic resonance imaging at 7 to 12 years of age.

Results: Three trajectory classes were identified among these youth: 53.50% of children exhibited elevated irritability during preschool that decreased longitudinally, 30.26% exhibited consistently low irritability, and 16.24% exhibited consistently elevated irritability. Compared with other classes, the elevated irritability class exhibited higher rates of maternal depression, early life adversity, later psychiatric diagnoses, and functional impairment. Further, elevated baseline irritability predicted later depression beyond adversity and personal and maternal depression history. The elevated irritability class exhibited a thicker cortex in the left superior frontal and temporal gyri and the right inferior parietal lobule.

Conclusion: Irritability manifested with specific developmental trajectories in this sample enriched for early depression. Persistently elevated irritability predicted poor psychiatric outcomes, higher risk for later depression, and decreased overall function later in development. Greater frontal, temporal, and parietal cortical thickness also was found, providing neural correlates of this risk trajectory.

Key words: irritability, development, latent trajectory, magnetic resonance imaging, cortical thickness

J Am Acad Child Adolesc Psychiatry 2018;57(5):336–342.

Associations between irritability and depression are of interest in developmental psychopathology. The unique role that irritability plays in pediatric depression is evident in that, in multiple editions of the *DSM*, irritability is a criterion for pediatric, but not adult, depressive disorders. Indeed, youth irritability predicts subsequent depression, and offspring of depressed mothers face increased risk for irritability and subsequent depression.[1-4] Moreover, phenomenological work has begun to elucidate distinct developmental trajectories of irritability in youth,[2,5] including associations with maternal depression. However, it is essential to integrate brain development into this research and examine longitudinal outcomes. We present the first study linking trajectories of irritability severity to brain structure differences in a pediatric sample enriched for depression. This adds to a small but growing literature examining the development of irritability, its neural underpinnings, and associations with depression and other clinical outcomes.

It is important to begin longitudinal studies of irritability and depression as early as preschool. Although temper outbursts are ubiquitous in preschoolers, normative and non-normative outbursts can be differentiated, providing opportunities for early interventions.[6-8] Early intervention could be particularly important in preschoolers also showing symptoms of depression and/or anxiety. The frequency of temper loss generally decreases during development, albeit with interindividual variability.[7] Prior population-based work identified several latent trajectories of irritability, in which most children showed low irritability that decreased across the first 9 years of life, but smaller subsamples exhibited chronically elevated or increasing irritability, which were more likely to have a maternal depression history and to develop more internalizing symptoms themselves.[2] Another longitudinal study similarly found a small group of children with chronically elevated irritability trajectories exhibiting the worst outcomes, including antisocial personality features in adulthood.[5] Similar trajectories have been noted across adolescence and young adulthood when examining irritability related to aggression and violence.[9] Critically, irritability relates to parental history of depression and anxiety and to risk for subsequent depression and impairment.[3,4] Indeed, across ages, meta-analysis longitudinally associates irritability with risk for depression, anxiety disorders, and global impairment,[1,10] whereas other studies highlight significant genetic associations between depression and irritability.[11,12]

Complementing these behavioral studies, recent work has probed the neural correlates of clinically impairing irritability in youth, although none in a pediatric sample with elevated early symptoms of depression.

Irritable youth show dysfunctional frontal, cingulate, and striatal responses to induced frustration[13,14] and aberrant frontal and amygdala responses during emotional face processing.[15-17] However, few studies have examined irritability-related differences in brain structure. Youth with severe irritability (specifically, disruptive mood dysregulation disorder) exhibit smaller gray matter volume in the pre-supplementary motor area,[18] dorsolateral prefrontal cortex,[18,19] superior frontal gyrus (SFG),[19] and larger insula volume[18] compared with healthy youth.

The present study leveraged a pediatric sample enriched for early depressive symptoms and followed prospectively to examine trajectories of irritability severity from preschool through late childhood/early adolescence. Importantly, this is the first study to examine differences in brain structure as a function of these trajectories; specifically, we measured cortical thickness across 3 neuroimaging waves during late childhood/early adolescence. Further, we add to the growing literature on irritability by focusing specifically on a sample enriched for early depression. Similar to prior work,[2] we expected classes of youth with low and declining or chronically elevated or increasing irritability trajectories across development. We hypothesized that classes with elevated irritability would show higher rates of depressive symptoms compared with others. Also, building on the affective disorders literature, we hypothesized associations between irritability and cortical thickness in prefrontal cortical regions involved in cognitive control and emotion regulation.

METHOD

Participants

Participants were drawn from the prospective longitudinal Preschool Depression Study (PDS; N = 306; present analyses examined 271 with $\geq$3 annual clinical assessments available as is necessary for growth curve modeling; Table S1, available online, presents demographics by subsample). The PDS has the broad goal of exploring clinical and neural outcomes relating to preschool-onset depression. Details of the study have been published previously.[20] Briefly, 3- to 5-year-old children and their primary caregivers were recruited from the St. Louis metropolitan area with oversampling for early depressive symptoms. Children and caregivers completed in-depth clinical interviews annually; children participated in 3 annual neuroimaging scans beginning at 7 to 12 years of age. Parental written consent and child assent after 4 years of age were obtained before participation and the institutional review board at Washington University (St. Louis, MO) approved all experimental procedures.

Diagnostic Assessments

Staff, trained to proficiency by appropriate experts, conducted in-person assessments with participants and parents/guardians from enrollment through the time of the scans (Figure S1, available online, presents study flow diagram). Before children were 8 years of age, a semistructured parent-report interview was used to assess psychiatric symptoms (Preschool-Age Psychiatric Assessment; PAPA).[21] After 8 years of age, the child and parent reports on the Childhood and Adolescent Psychiatric Assessment (CAPA)[22] were collected. Interviews were audiotaped, reviewed for reliability, and calibrated for accuracy.[20]

PAPA and CAPA data were summarized based on whether children ever received diagnoses of major depressive disorder (MDD), an anxiety disorder (generalized anxiety, separation anxiety, panic disorder, agoraphobia, and/or social phobia), oppositional defiant disorder (ODD) or conduct disorder (CD), or attention-deficit/hyperactivity disorder (ADHD) at any time point from baseline through the ninth annual assessment. Further, the presence or absence of 15 stressors endorsed at baseline was summed plus early exposure to poverty to create a composite early life adversity score (range 0–16; details are presented in Supplement 1, available online). Psychotropic medication use also was assessed, because this varied widely across participants and time; a summary variable of use ever was examined.

Measures

Irritability severity was defined as the sum of 5 items from the depression and conduct problems sections of the PAPA and CAPA (items: losing temper, nondestructive temper tantrums, irritability, touchy or easily annoyed, angry, or resentful). Supplement 1, available online, presents full definitions, stem questions, and coding rules for these items. Similar to prior work,[3] the presence of symptoms at each assessment was summed (range 0–5) based on PAPA parent report or if either reporter endorsed an item on the CAPA.[23] This measure was reliable across assessments (Cronbach $\alpha = 0.75$).

Parents rated children's functional impairment at the first scan session on the MacArthur Health and Behavior Questionnaire (HBQ).[24] Three HBQ subscales were of interest: functional impairment (range 0–16), global peer relationships (range 0–22), and academic functioning (range 0–32).

The Family Interview for Genetic Studies[25] assessed the presence of affective disorders in first- and second-degree relatives. The presence or absence of maternal depression or anxiety history was examined as a risk factor for internalizing symptoms in the child.

The Kaufman Brief Intelligence Test[26] or the Wechsler Abbreviated Scale of Intelligence[27] was used to assess IQ at school age (depending on age or study wave). Baseline socioeconomic status was quantified as a family's income-to-needs ratio[28,29] (i.e., total family income divided by the federal poverty level, based on family size, at the time of assessment).

Examining Trajectories of Irritability Across Development

We used latent class mixture models implemented in R[30] using the lcmm package[31,32] to identify latent groups of individuals who shared similar developmental trajectories of irritability severity. In these models, irritability was the dependent variable and age at each assessment (in months) was the independent variable. Details are presented in Supplement 1, available online.

Demographics and Clinical Outcomes

Baseline demographic differences (Table 1) across the 3 latent trajectory classes were characterized using analysis of variance and χ^2 test. The χ^2 tests characterized group differences in maternal depression history and child diagnoses (presence or absence across any of up to 8 assessments from baseline through approximately 15 years old). Analyses of variance assessed group differences in later functional impairment (HBQ).

Baseline Irritability as a Predictor

In addition, we used logistic regression to test whether baseline irritability predicted ever receiving a depression diagnosis above and beyond standard predictive/risk factors—sex, baseline PAPA depression symptom severity (excluding irritability, including negative mood, anhedonia, somatic changes, guilt, etc.), adversity, and maternal depression history. To examine specificity, we used a matched logistic regression with irritability predicting anxiety diagnosis, beyond sex, baseline anxiety severity (e.g.,

TABLE 1 Demographics and Clinical Characteristics of Latent Trajectory Classes

	Low Irritability	Declining Irritability	High Irritability
Class, n	82	145	44
Assessments (n), mean (SD)	6.01 (1.36)	6.14 (1.41)	5.70 (1.56)
Girls, n (%)***	41 (50.00)	81 (55.90)	10 (22.70)
White race, n (%)	40 (48.80)	81 (55.90)	28 (63.60)
Age at baseline (mo), mean (SD)***	53.64 (9.91)	51.72 (8.60)	59.07 (9.25)
Age at scan 1 (mo), mean (SD)	125.15 (13.81)	120.45 (14.84)	127.08 (15.56)
IQ, mean (SD)[a]	103.52 (13.63)	106.73 (14.89)	100.31 (15.47)
Baseline adversity score, mean (SD)*[b]	2.36 (1.66)	2.69 (1.79)	3.48 (2.40)
Baseline income-to-needs ratio, mean (SD)	2.31 (1.17)	2.19 (1.29)	2.24 (1.34)
HBQ peer relationships, mean (SD)*[c]	3.61 (0.59)	3.48 (0.50)	3.23 (0.77)
HBQ academic functioning, mean (SD)***[c]	4.22 (0.45)	3.98 (0.58)	3.34 (0.87)
HBQ global functioning, mean (SD)***[c]	1.96 (2.58)	4.05 (4.95)	9.67 (5.52)
Maternal MDD, n (%)***[d]	19 (23.20)	60 (41.40)	29 (69.00)
Child MDD, n (%)***	20 (24.40)	79 (54.50)	39 (88.60)
Child anxiety, n (%)***	37 (45.10)	83 (57.20)	39 (88.60)
Child ADHD, n (%)***	10 (12.20)	52 (35.90)	32 (72.70)
Child ODD/CD, n (%)***	11 (13.40)	57 (39.30)	35 (79.50)
Psychotropic medication use, n (%)***	4 (8.30)	22 (23.4)	16 (61.50)

*Note: The number of participants in each trajectory class (class) and the number of times that each child's symptomology was assessed (assessments) are noted. Demographic characteristics are noted by group and by early adversity, impairment on the MacArthur Health and Behavior Questionnaire (HBQ), maternal depression history, and rates of child diagnoses across assessments. Group differences in these factors were tested by analysis of variance or χ^2 test and significant group differences are indicated (*p < .05; ***p < .001). We note the total sample with available data on each measure and the sample size within each class. ADHD = attention-deficit/hyperactivity disorder; CD = conduct disorder; MDD = major depressive disorder; ODD = oppositional defiant disorder.*

[a]*n = 220 IQ, n = 65 low, n = 120 declining, n = 35 high irritability.*

[b]*n = 269 adversity, n = 80 low, n = 145 declining, n = 44 high irritability.*

[c]*n = 160 HBQ, n = 47 low, n = 89 declining, n = 24 high irritability.*

[d]*n = 269 Maternal MDD, n = 82 low, n = 145 declining, n = 42 high irritability.*

worrying, avoidance, fear/anxiety of separation, autonomic symptoms, etc.), adversity, and maternal anxiety history.

Magnetic Resonance Image Acquisition and Structural Analysis

Three waves of neuroimaging data were collected using a 3-T TIM TRIO Siemens scanner (Siemens, Erlangen, Germany). Structural images (1-mm isotropic voxels) were processed using the longitudinal stream[33,34] in FreeSurfer 5.3 (http://surfer.nmr.mgh.harvard.edu/). Details are presented in Supplement 1, available online.

Of the 271 children with estimated irritability trajectories, 139 had 2 to 3 waves of structural data usable for longitudinal processing. Figure S2, available online, shows the timing of these scans. Differences in cortical thickness were compared across classes controlling for age at the first scan and sex using a "different offset, different slope" general linear model in FreeSurfer (mri_glmfit). Specifically, 2 standard summary outcome measures from the longitudinal processing model were examined: temporal average (thickness at midpoint of linear fit) and rate of change in thickness across time. Cluster-wise correction was used for multiple comparisons (mri_glmfit-sim–cache 1.3 abs–cwp 0.05–2spaces) with a vertex-wise/cluster-forming threshold of a p value less than .05 and adjusting for tests in each hemisphere. Interactions with age and sex are presented in Table S2, available online, because we did not have a priori hypotheses about these interactions and power was limited, given the small samples split within class by sex. Exploratory analyses examining potential interactions with quadratic effects of age also are noted in Supplement 1, available online.

RESULTS

Irritability Trajectory Characteristics

Of the 5 models tested (Table S3, available online), the model of 3 latent classes (Figure 1) showed the lowest Bayesian information criterion, the lowest Akaike information criterion, and the highest entropy, suggesting the best and most parsimonious fit and best classification of individuals. Some children (14.24%) showed high irritability during the preschool period, which remained high across development ("high irritability" class). Many children (53.14%) showed preschool-age irritability as elevated as the high irritability class, but decreasing across development ("declining irritability" class). This likely represents a normative trajectory across development.[7] Some children (30.63%) showed low preschool-age irritability, which remained low ("low irritability" class).

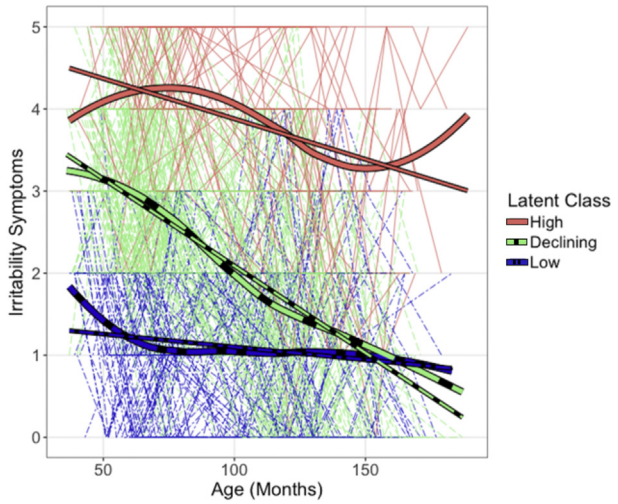
The children in the high irritability class were more likely to be male and slightly older at baseline and had more adverse childhood experiences (Table 1). The trajectory classes did not differ significantly in race, age at scan, IQ, or income-to-needs ratio.

Clinical Associations

Compared with the other classes, children in the high irritability class showed higher rates of maternal depression history (Table 1). These youths were more likely to receive a diagnosis of MDD, anxiety disorder, ADHD, or ODD/CD across clinical assessment waves (Table 1).

Baseline levels of irritability, which were elevated in the high irritability class, predicted later MDD diagnoses (odds ratio 1.37, $p = .01$) beyond the effects of sex, maternal depression, adversity, and other baseline depressive symptoms (Table S4, available online). Indeed, in area

FIGURE 1 Irritability Trajectory Classes



Note: The low, declining, and high irritability latent trajectory classes are displayed. Individual participant irritability severity scores (y-axis; Preschool-Age Psychiatric Assessment [PAPA] and Childhood and Adolescent Psychiatric Assessment [CAPA] sum scores) at each assessment wave are presented as thin lines (i.e., 1 line per participant) and time locked to the child's age at assessment indicated in months on the x-axis. Thicker lines represent the linear slope and loess fit lines of irritability severity for each trajectory class across age.

under the curve analyses predicting who did versus did not develop MDD, baseline irritability alone (area under the curve 0.73, 95% CI 0.68–0.79) was not significantly different in its discrimination from full baseline MDD symptom severity (area under the curve 0.82, 95% CI 0.77–0.87). The association between irritability and depression showed specificity in not predicting later anxiety diagnosis (odds ratio 1.02, $p = .89$; Table S4, available online).

Brain Structure

No regions showed significant differences in rate of change of thickness over time as a function of irritability class. Five regions did show overall differences in cortical thickness (temporal average, thickness at midpoint of linear fit) as a function of irritability trajectory classes (Table 2, Figure 2). The left SFG, left superior temporal gyrus (STG), and right inferior parietal lobule (IPL) showed thicker gray matter in children in the high irritability class; these regions showed the same pattern at each wave (Figure S3, available online). The left rostral middle frontal gyrus and right SFG showed a different pattern, specifically a thinner cortex in the declining irritability class compared with the other 2 classes.

Various control analyses were examined to confirm these findings. The number of children ever taking psychotropic medications by the time of the first scan differed by trajectory class (Table 1). However, all differences in cortical thickness across trajectory classes remained significant after controlling for medication use ($p < .05$). Of these 5 regions, only left SFG thickness differed significantly by medication use, showing a thinner cortex in those with prior medication use (Table S5, available online). Medication use was too varied across children and time to examine effects of specific medications. In addition, we found no significant differences in cortical thickness when comparing participants who had or had not ever been diagnosed with MDD, anxiety, ODD/CD, or ADHD (Table S5, available online), suggesting some specificity to irritability trajectories. These associations between trajectory groups also remained significant when controlling for baseline levels of non-irritability depression symptom severity, adverse life events, income-to-needs ratio, and irritability severity at scan. These covariates did not independently associate with thickness in any of these regions (Table S5, available online).

DISCUSSION

This is the first study to relate latent trajectories of irritability symptoms across early development to cortical thickness in a sample of youth enriched for depressive symptoms during the preschool period. We identified 3 latent trajectories classes based on repeated semistructured clinical interviews from preschool age through adolescence. The class with chronically elevated irritability showed poor clinical outcomes, specifically increased rates of depression and other psychopathology and greater general functional impairment. This trajectory class experienced higher rates of maternal depression and early adversity. Furthermore, during later childhood/early adolescence, these chronically irritable children showed greater SFG, STG, and IPL thickness across 3 scan waves compared with other children in the sample. This work builds on the small prior literature examining trajectories and longitudinal outcomes of irritability. In addition, this work presents novel findings in a sample with elevated early depressive symptoms and in regarding cortical thickness.

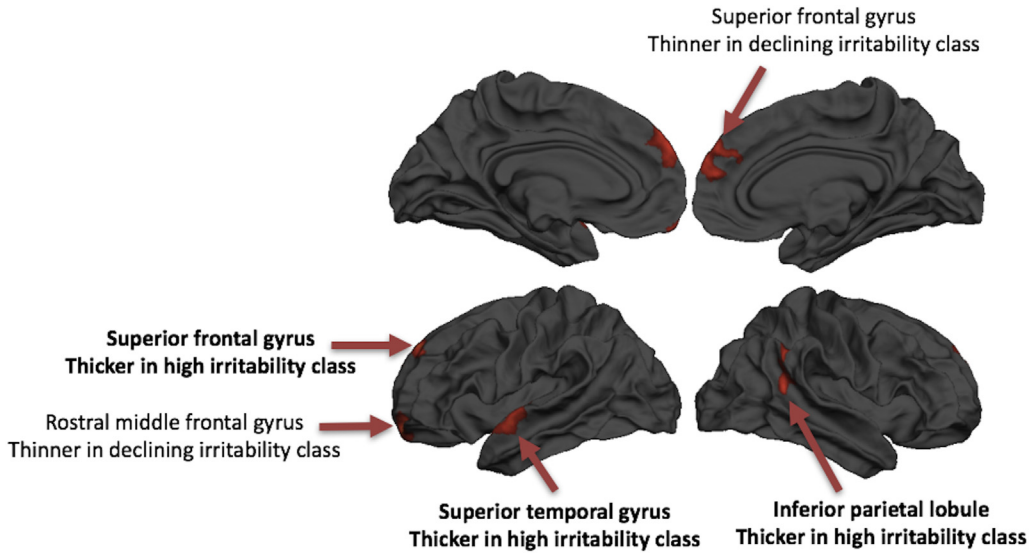
Although this sample was enriched for children with elevated early depression symptoms, the latent trajectory classes identified resemble prior

TABLE 2 Regions Showing Differences in Cortical Thickness by Trajectory Class

	Peak Coordinates						Estimated Marginal Means by Class			
Region	***x***	***y***	***z***	**Size (mm^2)**	***F***	**Partial η^2**	**L**	**D**	**H**	**Effect**
Left superior frontal gyrus	−10	48	33	468	9.09	0.13	3.26	3.24	3.46	L = D < H
Left superior temporal gyrus	−56	−4	−7	504	5.10	0.08	3.08	3.04	3.20	L = D < H
Right inferior parietal lobule	45	−48	15	450	8.84	0.12	2.74	2.70	2.87	L = D < H
Left rostral middle frontal gyrus	−16	53	−12	438	12.03	0.16	3.06	2.89	3.09	D < L = H
Right superior frontal gyrus	8	56	21	446	5.44	0.08	3.26	3.14	3.23	D < L = H

Note: Five regions showed significant trajectory class-related differences in the temporal average of cortical thickness across scan waves. The Talairach coordinates at the peak of thickness differences are indicated (x, y, z) as is the area of each cluster (size). The average thickness across each cluster was extracted for each participant; the F statistic and effect size (partial η^2) of group differences in mean thickness from a univariate general linear model controlling for age and sex are indicated. The estimated marginal means from these models are indicated for each trajectory class as is a summary of the Bonferroni-corrected significant ($p < .05$) post hoc differences between classes (effect). D = declining irritability; H = high irritability; L = low irritability.

FIGURE 2 Regions Showing Differences in Cortical Thickness by Trajectory Class



Note: Five regions showing differences in the temporal average of cortical thickness across wave as a function of irritability trajectory class are displayed. Regions showing greater thickness in the high irritability class are denoted in bold.

research. A large community-based study using parent-report questionnaire data from 1 to 9 years of age identified 5 latent trajectory classes.[2] The large sample in this prior study might have allowed for the identification of smaller trajectory classes that we did not detect. We did replicate findings that most children show a normative decline in irritability across early childhood,[2] although children in the present sample showed higher levels of irritability at baseline, possibly because of the high levels of early depressive symptoms. Further, we replicated a class of children exhibiting chronically high irritability who were more likely to have maternal depression history. We also noted a larger male proportion in this class, which could relate to our sample characteristics/acquisition but prior work did note a large male percentage in a high and steady irritability class.[2] Our work also identified similar but fewer classes than the work by Hawes *et al.*[5] Similarly, we found the worst psychiatric outcomes among children with early irritability that remained high across development, although this prior study found that most children were in a chronically low irritability class.[5] Notably, a strength and unique feature of the present study was the well-validated clinical interview assessment rather than questionnaire measures and the use of a preschool sample oversampled for depressive symptoms.

In examining cortical thickness in later childhood/early adolescence, we found that the left SFG, left STG, and right IPL showed increased thickness in the high irritability class compared with the other trajectory classes. This was stable across 3 neuroimaging waves. This also was robust when controlling for early depressive symptoms, baseline adversity, and baseline socioeconomic status, indicating some specific to irritability in general or in the context of elevated depressive symptoms. Prior work identified reduced gray matter volume in a similar but right hemisphere SFG region in children with disruptive mood dysregulation disorder, bipolar disorder, or ADHD compared with healthy youth.[19] However, ours is the first study to link developmental trajectories of irritability, based on a dimensional measure, to brain structure, and the first to examine this in a sample enriched for early onset depression. We also identified 2 regions showing a thinner cortex in the declining irritability class compared with the other classes, which was not hypothesized and should be explored in future work. In addition, future work is needed to examine the functional implications of these alterations in SFG, STG, and IPL structure. Given documented neuropsychological aberrations in pediatric irritability (e.g.,[35-38]), these structural alterations could relate to deficits in attentional or executive control and potentially to reward processing or social cognition. In particular, prior work highlighted a similar SFG region as showing altered reward/loss processing in irritable children[13] and altered response to negative feedback in the IPL.[14]

Given the limited number of brain morphometry studies on youth with elevated irritability or on psychiatrically impaired children scanned longitudinally, it is difficult to draw conclusions regarding increased cortical thickness in chronic irritability. Prior cross-sectional research on pediatric psychopathology showed mixed results. Studies of pediatric anxiety disorders[39-42] and depression[43] suggested increased regional thickness or gray matter volume. Given cross-sectional, longitudinal, and genetic associations among irritability, anxiety, and depression,[1,10] increased cortical thickness could be a shared feature among the 3 symptoms. Other studies suggested a thinner cortex associated with ADHD,[44,45] CD,[46,47] and anxiety, although the latter study highlighted weaker associations in older youth.[48] A study also suggested slower patterns of typical growth and reduction of different subcortical volumes in youth with depression.[49] Longitudinal imaging studies suggested that youth with ADHD reach peak cortical thickness later than controls[50] and that greater cortical thinning is associated with worse ADHD outcomes.[51] The timing of neuroimaging across development could contribute to these mixed results. When interpreting increased cortical thickness, it is important to note the typical developmental thinning in the regions identified here in this age range, although critically, we do not identify any group-by-time interactions. Thus, further longitudinal studies of irritability in samples with and without early depression and other pediatric psychiatric problems are needed to determine the clinical significance and implications of increased and decreased cortical thickness.

The present study informs the demographic and clinical characteristics of children with high irritability that remains stable across development. Of note, although many children exhibit irritability, only a subset manifests persistent irritability across development even in a

sample enriched for early depressive symptoms. This distinct, relatively small subset is worthy of clinical attention given the higher risk for later depression and general impairment. Importantly, elevated preschool-age irritability, particularly when persistent, might predict later depressive episodes and poorer clinical outcome over and above risk owing to familial depression history and early adversity. Our findings suggest the need to carefully monitor children with irritability, especially those who also have depressive symptoms, and to evaluate familial factors, such as maternal depression and early adversity. The present data suggest that children with persistently elevated irritability might be most in need of early intervention and prevention given the elevated risks of poor outcomes attributable to this potentially modifiable characteristic.

The present study replicates and provides further robust evidence for increased maternal depression and child psychopathology among children with chronically high irritability.[2] Because this sample was enriched for early depressive symptoms, the rates of maternal depression and child pathology are above the expected population rates. Nonetheless, the graded increase in risk and manifest illness across the low, declining, and high irritability trajectory groups is marked, mirroring prior findings.[2,3] Children in the high irritability class also experienced more early life adversity. Furthermore, we note that early irritability symptoms, which are particularly elevated in the high irritability group, predict later depression above and beyond other risk factors, including maternal depression, adversity, sex, and perhaps most notably even other early depressive symptoms. These data suggest that early irritability might be a robust risk marker and that monitoring these symptoms across development aid prediction of later depression. Future work should continue to confirm the trajectories of irritability across development to highlight additional risk factors and clinical outcomes. Brain imaging beginning earlier in development also will be critical to identifying potential risk markers for sustained irritability and to better understand associated neural development. It also will be important to parse correlates of these neural differences to identify particular behaviors or outcomes associate with regional structure.

Although the present approach to retrospectively ascertaining irritability severity from diagnostic interview has been used previously (e.g.,[52]) and similar approaches have been used with questionnaire measures (e.g.,[2,5]), future work should use measures specifically designed to assess irritability (e.g., the Affective Reactivity Index[53] or the Multidimensional Assessment of Preschool Disruptive Behavior[54]). These measures could yield more robust assessment of irritability across development and avoid biased recall. In addition, a possible limitation of this study could be an omitted variable bias in our analysis and the creation of trajectories. As such, the present work should be compared with future studies using prospective and targeted measures to assess the generalizability of these results. Further, although the well-validated clinical interview assessments were a strength of the present study, we used only parent report at the early assessment time points (child report was not obtained before 8 years of age) and then combined parent and child reports at later waves. This maximized our use of the available data and should not contribute to systematic differences among trajectory groups. In addition, although the present study examined 3 annual waves of high-quality neuroimaging, we did not have the data to assess brain structure early in development; earlier scanning could aid in predicting longitudinal outcomes and is an important future step. Moreover, although the oversampling for depressive symptoms in this sample was useful in examining associations with psychopathology, this might limit generalizability to population samples and thus future work in epidemiologic samples will be needed to replicate these effects.

In addition, psychotropic medication use differed across trajectory groups, but this did not affect the observed associations between irritability and cortical thickness. Medication use was related on its own to SFG thickness, although we were limited in our ability to parse effects of particular medications. Similarly, we observed differences in IQ between trajectory groups; that is, compared with the high irritability group, mean IQ was 6.42 points higher in the declining irritability group and 3.21 points higher in the low irritability group. Although these differences did not reach statistical significance and did not correlate with thickness in the regions identified here, we lacked the sample size to test moderation effects by IQ. Given associations between IQ and brain structure,[55] it will be important to examine this in future studies, particularly to explore interactions between IQ and irritability on brain development and depression risk. Although the overall sample was relatively large for neuroimaging studies, the subsamples in each trajectory group were too small to assess age and sex interactions.

Overall, in this sample enriched for early depression, we found evidence for distinct developmental trajectories of irritability across childhood and a unique role for irritability in brain development. Of note, irritability longitudinally predicted later depression in youth above and beyond typical risk factors in this enriched sample. Future studies with larger samples in community samples should examine these phenomena further. In particular, it will be interesting to explore potential moderation effects of age, sex, and other factors related to brain structure, such as psychotropic medication use, socioeconomic status,[29] IQ,[55] or cognitive control.[5]

Accepted March 9, 2018.

Dr. Pagliaccio is with the New York State Psychiatric Institute and College of Physicians and Surgeons, Columbia University, New York. Drs. Pine and Leibenluft are with the Emotion and Development Branch, National Institute of Mental Health (NIMH), Bethesda, MD. Drs. Barch and Luby are with the Washington University School of Medicine, St. Louis, MO.

This work was supported by NIMH grants R01MH66031, R01MH084840, R01MH090786, R01MH098454-S, and R01MH064769-06A1. Work by Drs. Pagliaccio, Pine, and Leibenluft was supported by the Intramural Research Program at the NIMH.

The authors thank all participants and their families who provided time and effort to make this study possible.

Disclosure: Drs. Pagliaccio, Pine, Barch, Luby, and Leibenluft report no biomedical financial interests or potential conflicts of interest.

Correspondence to: David Pagliaccio, PhD, Division of Child and Adolescent Psychiatry, Department of Psychiatry, New York State Psychiatric Institute and College of Physicians and Surgeons, Columbia University, 1051 Riverside Drive, New York, NY 10032; e-mail: david.pagliaccio@nyspi.columbia.edu

https://doi.org/10.1016/j.jaac.2018.02.010

REFERENCES

1. Vidal-Ribas P, Brotman MA, Valdivieso I, Leibenluft E, Stringaris A. The status of irritability in psychiatry: a conceptual and quantitative review. J Am Acad Child Adolesc Psychiatry. 2016;55:556-570.

2. Wiggins JL, Mitchell C, Stringaris A, Leibenluft E. Developmental trajectories of irritability and bidirectional associations with maternal depression. J Am Acad Child Adolesc Psychiatry. 2014;53:1191-1205.

3. Dougherty LR, Smith VC, Bufferd SJ, *et al*. Preschool irritability: longitudinal associations with psychiatric disorders at age 6 and parental psychopathology. J Am Acad Child Adolesc Psychiatry. 2013;52:1304-1313.
4. Whelan YM, Leibenluft E, Stringaris A, Barker ED. Pathways from maternal depressive symptoms to adolescent depressive symptoms: the unique contribution of irritability symptoms. J Child Psychol Psychiatry. 2015;56:1092-1100.
5. Hawes SW, Perlman SB, Byrd AL, Raine A, Loeber R, Pardini DA. Chronic anger as a precursor to adult antisocial personality features: the moderating influence of cognitive control. J Abnorm Psychol. 2016;125:64-74.
6. Keenan K, Wakschlag LS. More than the terrible twos: the nature and severity of behavior problems in clinic-referred preschool children. J Abnorm Child Psychol. 2000;28:33-46.
7. Wakschlag LS, Estabrook R, Petitclerc A, *et al*. Clinical implications of a dimensional approach: the normal:abnormal spectrum of early irritability. J Am Acad Child Adolesc Psychiatry. 2015;54:626-634.
8. Belden AC, Thomson NR, Luby JL. Temper tantrums in healthy versus depressed and disruptive preschoolers: defining tantrum behaviors associated with clinical problems. J Pediatr. 2008;152:117-122.
9. Caprara GV, Paciello M, Gerbino M, Cugini C. Individual differences conducive to aggression and violence: trajectories and correlates of irritability and hostile rumination through adolescence. Aggress Behav. 2007;33:359-374.
10. Rice F, Sellers R, Hammerton G, *et al*. Antecedents of new-onset major depressive disorder in children and adolescents at high familial risk. JAMA Psychiatry. 2017;74:153-160.
11. Roberson-Nay R, Leibenluft E, Brotman MA, *et al*. Longitudinal stability of genetic and environmental influences on irritability: from childhood to young adulthood. Am J Psychiatry. 2015;172:657-664.
12. Stringaris A, Zavos H, Leibenluft E, Maughan B, Eley TC. Adolescent irritability: phenotypic associations and genetic links with depressed mood. Am J Psychiatry. 2012; 169:47-54.
13. Perlman SB, Jones BM, Wakschlag LS, Axelson D, Birmaher B, Phillips ML. Neural substrates of child irritability in typically developing and psychiatric populations. Dev Cogn Neurosci. 2015;14:71-80.
14. Deveney CM, Connolly ME, Haring CT, *et al*. Neural mechanisms of frustration in chronically irritable children. Am J Psychiatry. 2013;170:1186-1194.
15. Brotman MA, Rich BA, Guyer AE, *et al*. Amygdala activation during emotion processing of neutral faces in children with severe mood dysregulation versus ADHD or bipolar disorder. Am J Psychiatry. 2010;167:61-69.
16. Stoddard J, Tseng WL, Kim P, *et al*. Association of irritability and anxiety with the neural mechanisms of implicit face emotion processing in youths with psychopathology. JAMA Psychiatry. 2017;74:95-103.
17. Wiggins JL, Brotman MA, Adleman NE, *et al*. Neural correlates of irritability in disruptive mood dysregulation and bipolar disorders. Am J Psychiatry. 2016;173:722-730.
18. Adleman NE, Fromm SJ, Razdan V, *et al*. Cross-sectional and longitudinal abnormalities in brain structure in children with severe mood dysregulation or bipolar disorder. J Child Psychol Psychiatry. 2012;53:1149-1156.
19. Gold AL, Brotman MA, Adleman NE, *et al*. Comparing brain morphometry across multiple childhood psychiatric disorders. J Am Acad Child Adolesc Psychiatry. 2016;55:1027-1037.
20. Luby JL, Si X, Belden AC, Tandon M, Spitznagel E. Preschool depression: homotypic continuity and course over 24 months. Arch Gen Psychiatry. 2009;66:897-905.
21. Egger HL, Ascher B, Angold A. The Preschool Age Psychiatric Assessment: Version 1.4. Durham, NC: Center for Developmental Epidemiology, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center; 2003.
22. Angold A, Costello EJ. The Child and Adolescent Psychiatric Assessment (CAPA). J Am Acad Child Adolesc Psychiatry. 2000;39:39-48.
23. Bird HR, Gould MS, Staghezza B. Aggregating data from multiple informants in child psychiatry epidemiological research. J Am Acad Child Adolesc Psychiatry. 1992;31:78-85.
24. Armstrong JM, Goldstein LH; The MacArthur Working Group on Outcome Assessment. Manual for the MacArthur Health and Behavior Questionnaire (HBQ 1.0). Pittsburgh: MacArthur Foundation Research Network on Psychopathology and Development (David J. Kupfer, Chair), University of Pittsburgh; 2003.
25. Maxwell ME. Family Interview for Genetic Studies (FIGS): A Manual for FIGS. Bethesda, MD: Clinical Neurogenetics Branch, Intramural Research Program, National Institute of Mental Health; 1992.
26. Kaufman AS, Kaufman NL. Kaufman Brief Intelligence Test. Hoboken, NJ: Wiley Online Library; 2004.
27. Wechsler D. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation; 1999.
28. Barch D, Pagliaccio D, Belden A, *et al*. Effect of hippocampal and amygdala connectivity on the relationship between preschool poverty and school-age depression. Am J Psychiatry. 2016;173:625-634.
29. Luby J, Belden A, Botteron K, *et al*. The effects of poverty on childhood brain development: the mediating effect of caregiving and stressful life events. JAMA Pediatr. 2013; 167:1135-1142.
30. R: A Language and Environment for Statistical Computing [computer program]. Vienna, Austria: R Foundation for Statistical Computing; 2015.
31. Proust-Lima C, Philipps V, Diakite A, Liquet B, Proust MC. Package 'lcmm'. https://cran.r-project.org/web/packages/lcmm/lcmm.pdf. Published 2016. Accessed April 13, 2018.
32. Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using latent classes and latent processes: the R package lcmm. J Stat Softw. 2017;78:1-56 (arXiv: 1503.00890 [stat.CO]).
33. Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage. 2012;61:1402-1418.
34. Reuter M, Fischl B. Avoiding asymmetry-induced bias in longitudinal image processing. Neuroimage. 2011;57:19-21.
35. Salum GA, Mogg K, Bradley BP, *et al*. Association between irritability and bias in attention orienting to threat in children and adolescents. J Child Psychol Psychiatry. 2017;58:595-602.
36. Rich BA, Brotman MA, Dickstein DP, Mitchell DG, Blair RJ, Leibenluft E. Deficits in attention to emotional stimuli distinguish youth with severe mood dysregulation from youth with bipolar disorder. J Abnorm Child Psychol. 2010;38:695-706.
37. Adleman NE, Kayser R, Dickstein D, Blair RJ, Pine D, Leibenluft E. Neural correlates of reversal learning in severe mood dysregulation and pediatric bipolar disorder. J Am Acad Child Adolesc Psychiatry. 2011;50:1173-1185.e1172.
38. Pagliaccio D, Wiggins JL, Adleman NE, *et al*. Behavioral and neural sustained attention deficits in disruptive mood dysregulation disorder and attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2017;56:426-435.
39. Strawn JR, Wehry AM, Chu WJ, *et al*. Neuroanatomic abnormalities in adolescents with generalized anxiety disorder: a voxel-based morphometry study. Depress Anxiety. 2013; 30:842-848.
40. Gold AL, Steuber ER, White LK, *et al*. Cortical thickness and subcortical gray matter volume in pediatric anxiety disorders. Neuropsychopharmacology. 2017;42:2423-2433.
41. Strawn JR, Hamm L, Fitzgerald DA, Fitzgerald KD, Monk CS, Phan KL. Neurostructural abnormalities in pediatric anxiety disorders. J Anxiety Disord. 2015;32:81-88.
42. Strawn JR, John Wegman C, Dominick KC, *et al*. Cortical surface anatomy in pediatric patients with generalized anxiety disorder. J Anxiety Disord. 2014;28:717-723.
43. Reynolds S, Carrey N, Jaworska N, Langevin LM, Yang XR, Macmaster FP. Cortical thickness in youth with major depressive disorder. BMC Psychiatry. 2014;14:83.
44. Almeida Montes LG, Prado Alcantara H, Martinez Garcia RB, De La Torre LB, Avila Acosta D, Duarte MG. Brain cortical thickness in ADHD: age, sex, and clinical correlations. J Atten Disord. 2013;17:641-654.
45. Almeida LG, Ricardo-Garcell J, Prado H, *et al*. Reduced right frontal cortical thickness in children, adolescents and adults with ADHD and its correlation to clinical variables: a cross-sectional study. J Psychiatr Res. 2010;44(16):1214-1223.
46. Hyatt CJ, Haney-Caron E, Stevens MC. Cortical thickness and folding deficits in conduct-disordered adolescents. Biol Psychiatry. 2012;72:207-214.
47. Wallace GL, White SF, Robustelli B, *et al*. Cortical and subcortical abnormalities in youths with conduct disorder and elevated callous-unemotional traits. J Am Acad Child Adolesc Psychiatry. 2014;53:456-465.e451.
48. Newman E, Thompson WK, Bartsch H, *et al*. Anxiety is related to indices of cortical maturation in typically developing children and adolescents. Brain Struct Funct. 2016; 221:3013-3025.
49. Whittle S, Lichter R, Dennison M, *et al*. Structural brain development and depression onset during adolescence: a prospective longitudinal study. Am J Psychiatry. 2014;171:564-571.
50. Shaw P, Eckstrand K, Sharp W, *et al*. Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. PNAS. 2007;104:19649-19654.
51. Shaw P, Malek M, Watson B, Greenstein D, de Rossi P, Sharp W. Trajectories of cerebral cortical development in childhood and adolescence and adult attention-deficit/hyperactivity disorder. Biol Psychiatry. 2013;74:599-606.
52. Copeland WE, Brotman MA, Costello EJ. Normative irritability in youth: developmental findings from the Great Smoky Mountains Study. J Am Acad Child Adolesc Psychiatry. 2015;54:635-642.
53. Stringaris A, Goodman R, Ferdinando S, *et al*. The Affective Reactivity Index: a concise irritability scale for clinical and research settings. J Child Psychol Psychiatry. 2012;53: 1109-1117.
54. Wakschlag LS, Briggs-Gowan MJ, Choi SW, *et al*. Advancing a multidimensional, developmental spectrum approach to preschool disruptive behavior. J Am Acad Child Adolesc Psychiatry. 2014;53:82-96.e83.
55. Menary K, Collins PF, Porter JN, *et al*. Associations between cortical thickness and general intelligence in children, adolescents and young adults. Intelligence. 2013;41:597-606.

SUPPLEMENT 1

Adverse Childhood Experiences

1. poverty
2. parental arrests
3. parental hospitalizations
4. accident or crash with automobile, plane, or boat
5. accidental burning, poisoning, or drowning
6. attacked by an animal
7. adult loved one died
8. sibling or peer died
9. domestic violence
10. hospitalized, visited emergency room, or invasive medical procedure
11. man-made disaster
12. natural disaster
13. other life event
14. physical abuse
15. sexual abuse, sexual assault, or rape
16. witnessed someone threatened with harm, seriously injured, or killed

PAPA/CAPA Irritability Item Text

PGE0I01 Losing Temper. Discrete episodes of temper manifested by shouting or name calling but without violence and not meeting criteria for a temper tantrum.

- What sort of temper has s/he got?
- What happens when s/he loses his/her temper?

0 = absent
2 = present

PGE1I01 Nondestructive Temper Tantrums. Discrete episodes of excessive temper, frustration, or upset manifested by shouting, crying or stamping, or nondestructive violence directed against property.

Violence or damage done in this context does not constitute vandalism or assault.

- What sort of temper has s/he had in the past 3 months?
- What happens when something upsets him/her or s/he doesn't get what s/he wants?
- Does s/he have angry outbursts?
- Does s/he have temper tantrums?

0 = absent
2 = excessive temper, upset, shouting, crying, or nondestructive violence directed only against property (e.g., stamping, kicking, throwing toys, hitting walls, spitting, holding breath, etc.).

PDA8I01 Irritability. Increased ease of precipitation of externally directed feelings of anger, bad temper, short temper, resentment, or annoyance. (Change could predate the primary period and continue into at least part of the primary period.)

Note that this rating is of a change in the child's usual liability to be precipitated into anger; it does not refer to the form of the anger once it has been precipitated.

Note bene: The irritable mood itself is being rated, not just its manifestations; thus, frequency and duration ratings refer to the number and length of episodes of the mood, not of the episodes of snappiness, shouting, or quarrelsomeness.

- Has s/he been more irritable than usual in the last 3 months?
- Or made angry more easily?
- Has s/he had more tantrums than usual in the last 3 months?

0 = absent
2 = irritable mood present during at least 2 activities manifested by at least 1 instance of snappiness, shouting, or quarrelsomeness and at least sometimes uncontrollable by the child
3 = irritable mood present in most activities, accompanied by snappiness, shouting, or quarrelsomeness and nearly always uncontrollable by the child

PDA6I01 Touchy or Easily Annoyed. The child is generally more prone to FEELINGS of anger bad temper, short temper, resentment, sulking, or annoyance UNDER MINOR PROVOCATION than most children. This pattern need not represent a change in behavior.

- Do things get on his/her nerves easily?
- What sorts of things?
- Does s/he get annoyed more easily than most children, do you think?

0 = absent
2 = present

PDA7I01 Angry or Resentful. The child is generally more prone to MANIFESTATIONS of anger or resentment (such as snappiness, shouting, quarreling, or sulking) under minor provocation than most children.

This pattern need not represent a change in behavior.

- Does s/he get angry very often?
- Does s/he get "sulky" or "pout"?

0 = absent
2 = present

Examining Trajectories of Irritability Across Development

In the latent class mixture models, irritability was the dependent variable and age at each assessment (in months) was the independent variable. Age was indicated to have a random and class-specific effect. A linear link function was used and the variance-covariance matrix for the random effect was diagonal.

An automatic grid search was used to run the estimation function for a maximum of 20 iterations from 50 random sets of initial values to avoid convergence on local maxima. At the end, the parameters corresponding to the best logarithmic likelihood were used as initial values for the final estimation of the parameters.

Magnetic Resonance Image Acquisition and Structural Analysis

Three waves of neuroimaging data were collected using a 3-T TIM TRIO Siemens scanner. At each wave, 2 T1-weighted images were acquired in the sagittal plane using an MPRAGE 3-dimensional sequence (repetition time 2,400 ms, echo time 3.16 ms, flip angle 8°, slab 176 mm, 176 slices, matrix size 256 × 256, field of view 256 mm, voxel size 1 × 1 × 1 mm). The T1 image with the best image quality was used for structural analyses.

Structural images were processed using the longitudinal stream in FreeSurfer 5.3 (http://surfer.nmr.mgh.harvard.edu/).[1] Processing included skull stripping, Talairach transformations, atlas registration, and creating spherical surface maps and parcellations, initialized with common information from an unbiased within-patient template. This longitudinal stream decreases biases that could occur by selecting a single image as a baseline for registration and increases reliability and statistical power significantly.[2] White and pial surfaces generated by FreeSurfer were visually inspected by a single trained and supervised staff member. Surface edits were performed as needed and the surfaces were regenerated as recommended by FreeSurfer for quality control. From the overall Preschool Depression Study imaging sample, approximately 10% of scans had to be discarded for poor scan quality. These data were resampled on an average template (fsaverage) and smoothed with a 10-mm full-width/half-maximum kernel.

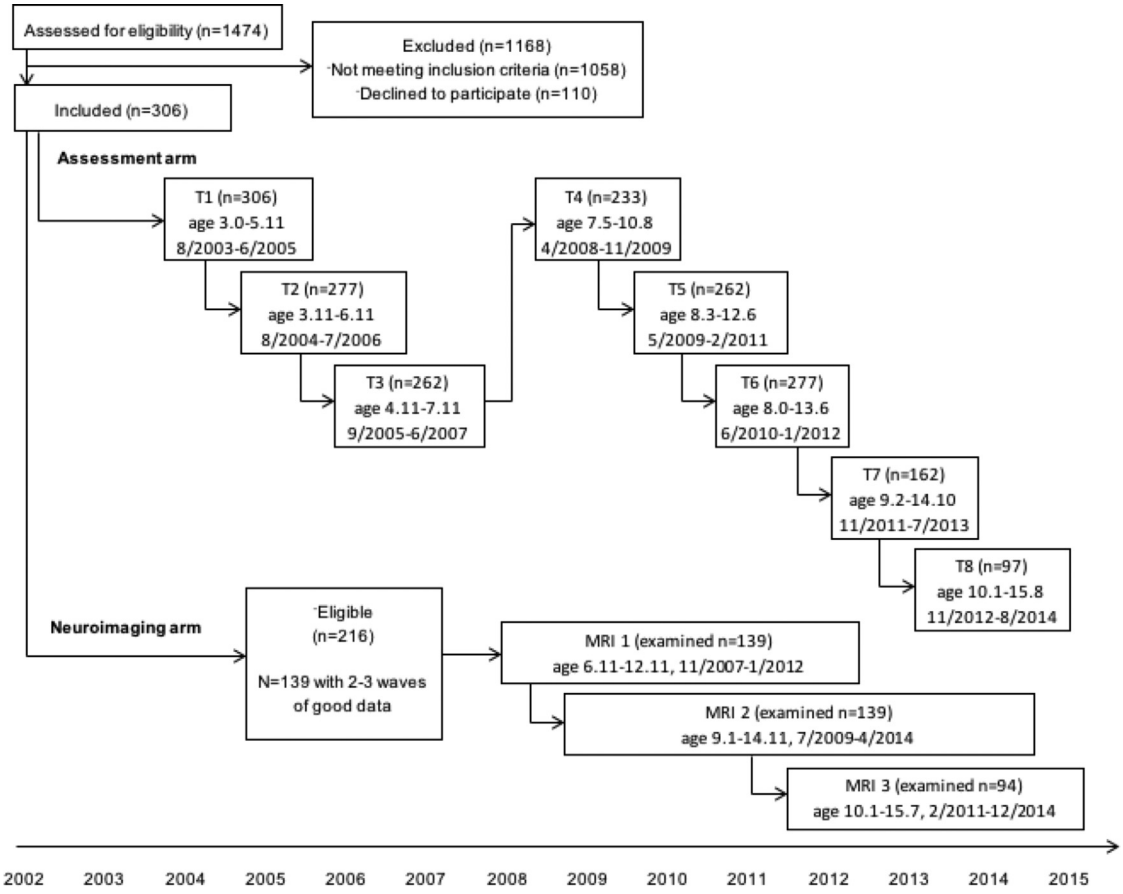
Testing Interactions with Quadratic Effects of Age

Because thickness can change nonlinearly with age, we examined a linear mixed model approach in Freesurfer[3,4] in addition to the analyses noted in the main text. In this model, we included a random intercept for each participant, trajectory group as a between-subject factor, and interactions with linear and quadratic age trends across waves. Quadratic interactions between age and trajectory group were examined to test for potential nonlinear differences in thickness. No regions passed multiple comparisons correction when examining class-by-quadratic age trend interactions in these linear mixed models.

SUPPLEMENTAL REFERENCES

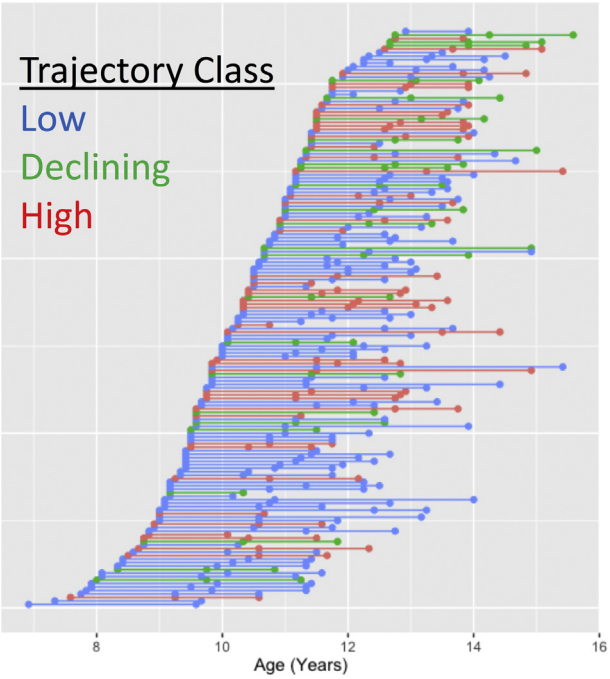
1. Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage. 2012;61:1402-1418.
2. Reuter M, Fischl B. Avoiding asymmetry-induced bias in longitudinal image processing. Neuroimage. 2011;57:19-21.
3. Bernal-Rusiel JL, Reuter M, Greve DN, Fischl B, Sabuncu MR. Alzheimer's Disease Neuroimaging I (2013): spatiotemporal linear mixed effects modeling for the mass-univariate analysis of longitudinal neuroimage data. Neuroimage. 2013;81:358-370.
4. Bernal-Rusiel JL, Greve DN, Reuter M, Fischl B, Sabuncu MR. Alzheimer's Disease Neuroimaging I (2013): statistical analysis of longitudinal neuroimage data with linear mixed effects models. Neuroimage. 2013;66:249-260.

FIGURE S1 Study Flow Diagram



Note: This figure displays the flow of data through the Preschool Depression Study. The total number of children through each assessment wave is presented. The number of children examined at each wave of imaging is presented. Only children with at least 2 waves of good-quality data were examined. MRI = magnetic resonance imaging.

FIGURE S2 Timing of Each Scan Wave With Good Structural Imaging Data



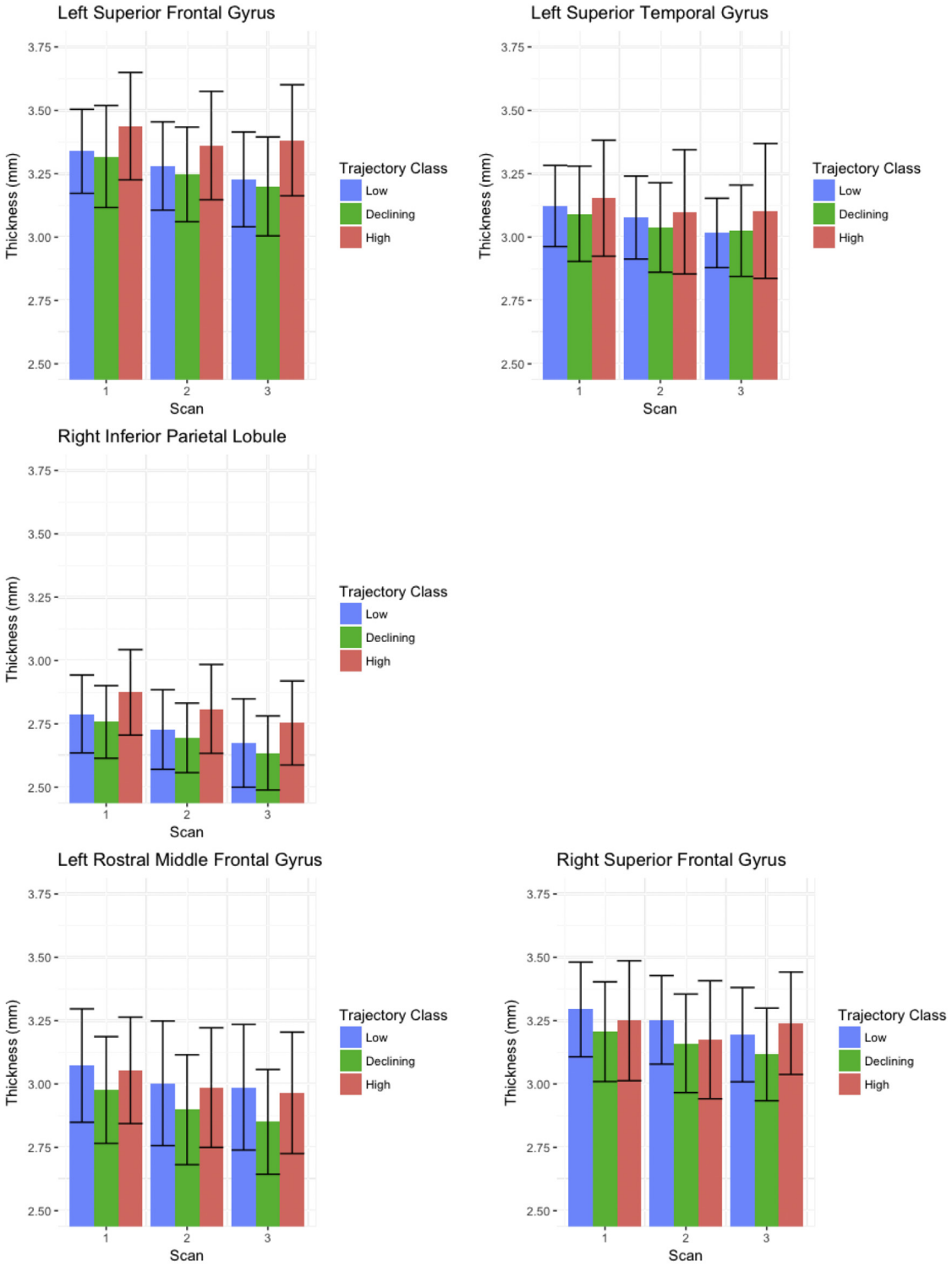
Note: Points connect each individual by lines. Individuals are sorted by age at the first scan and colored by irritability trajectory class. The first and second waves were collected on average 18.67 months apart (SD 6.52) and the second and third were similarly spaced (mean 15.53, SD 6.14).

TABLE S1 Demographic and Clinical Characteristics of Full Preschool Depression Study (PDS) Sample (N = 306), Sample of Children Included in the Latent Class Mixture Models (n = 271 Children With ≥3 Waves of Preschool-Age Psychiatric Assessment [PAPA] Data Available), and Subset of Those Children Included in Imaging Analyses (n = 139 Children with ≥2 Waves of Imaging Data)

	Original PDS Sample	Latent Class Subsample	Imaging Subsample
Class, n	306	271	139
Assessments (n), mean (SD)	5.5 (2.0)	6.0 (1.4)***	6.9 (0.9)*
Girls, n (%)	148 (48.4)	132 (48.7)	67 (48.2)
White race, n (%)	164 (53.6)	149 (55.0)	70 (50.4)
Age at baseline (mo), mean (SD)	53.4 (9.6)	53.5 (9.5)	54.4 (9.3)
Age at scan 1 (mo), mean (SD)	122.8 (14.8)	122.8 (14.8)	123.3 (14.8)
IQ, mean (SD)	104.7 (14.7)	104.8 (14.8)	104.5 (15.2)
Baseline adversity score, mean (SD)	2.7 (1.9)	2.7 (1.9)	2.9 (2.0)
Baseline income-to-needs ratio, mean (SD)	2.2 (1.3)	2.2 (1.3)*	2.0 (2.1)*
HBQ peer relationships, mean (SD)	3.4 (0.6)	3.4 (0.6)	3.4 (0.6)
HBQ academic functioning, mean (SD)	3.9 (0.7)	3.9 (0.7)	3.9 (0.7)
HBQ global functioning, mean (SD)	4.3 (5.1)	4.3 (5.1)	4.4 (5.2)
Maternal MDD, n (%)	119 (38.9)	108 (39.9)	65 (46.8)*
Child MDD, n (%)	152 (49.7)	138 (50.9)	82 (88.60)**
Child anxiety, n (%)	161 (52.6)	94 (34.7)*	53 (38.1)
Child ADHD, n (%)	107 (35.0)	152 (56.1)	82 (59.0)
Child ODD/CD, n (%)	133 (43.5)	117 (43.2)	63 (45.3)

*Note: Significant differences between those retained and excluded from each subsample are denoted in the latent class or imaging column (*p < .05; **p < .01; ***p < .001). Specifically, χ^2 and t tests were used to compare children included in the latent class subsample with those excluded from the original PDS sample (i.e., who did not have sufficient PAPA data). Compared with those excluded from the latent class analyses, the included children had more PAPA assessments (based on their inclusion criterion), tended to have higher income-to-needs ratios, and were more likely to have anxiety. Compared with those included in the latent class but excluded from the imaging analyses, the final 139 children in the imaging analyses completed more assessments, reported lower income-to-needs ratios, and exhibited more maternal and child depression diagnoses. ADHD = attention-deficit/hyperactivity disorder; CD = conduct disorder; MDD = major depressive disorder; ODD = oppositional defiant disorder.*

FIGURE S3 Mean Cortical Thickness From Each Cluster Showing Significant Whole Brain Trajectory Group Differences in the Temporal Average of Thickness, Split by Trajectory Group and Scan Wave



Note: *Error bars indicate 1 SD.*

TABLE S2 Regions Showing Interactions With Trajectory Class

Region	Peak Coordinates x	y	z	Size (mm^2)	F	Predicting	Interaction
Right superior frontal gyrus	9	53	11	466	3.36	average	class × sex
Left superior parietal lobule	−15	−82	27	428	3.15	average	class × sex
Left rostral middle frontal gyrus	−41	27	31	644	3.45	rate	class × age
Right posterior cingulate gyrus	9	−26	30	544	4.28	rate	class × age
Right precentral gyrus	11	−17	68	1013	3.80	rate	class × age
Right cingulate isthmus	22	−53	6	552	4.36	rate	class × age × sex
Right middle temporal gyrus	48	0	−32	1159	4.25	rate	class × age × sex
Right postcentral gyrus	52	−13	17	809	4.14	rate	class × age × sex

Note: These regions showed interactions between trajectory class and/or age and sex (interaction); the F statistic of this interaction is presented as are the peak coordinates and size of each cluster. These effects were significant in models examining the temporal average of cortical thickness or rate of change (predicting). The trajectory classes split by sex or age were small, so we did not explore these interactions because they were likely underpowered.

TABLE S3 Summary of Latent Class Mixture Models With 1 Through 5 Classes That Were Tested

Classes	Log Likelihood	Parameters	BIC	AIC	Entropy	Class 1, %	Class 2, %	Class 3, %	Class 4, %	Class 5, %
1	−2,685.99	5	5,400.00	5,129.09	—	100.00				
2	−2,672.84	8	5,390.49	5,116.79	0.39	69.00	31.00			
3	**−2,655.31**	**11**	**5,372.25**	**5,111.86**	**0.68**	**16.24**	**30.63**	**53.14**		
4	−2,651.06	14	5,380.56	5,118.89	0.65	7.75	29.52	23.62	39.11	
5	−2,644.71	17	5,384.66	5,117.89	0.67	8.86	26.20	38.75	15.13	11.07

Note: The result with 3 trajectory classes was used for subsequent analyses because it showed the lowest Bayesian information criterion (BIC), the lowest Akaike information criterion (AIC), and the highest entropy. Although models with 4 or 5 classes showed slightly lower log-likelihood values, the BIC and AIC values were higher, the entropy was slightly lower, and both had small trajectory classes containing less than 10% of the sample. Bold typeface denotes the model that was selected for the analyses in the current study.

TABLE S4 Logistic Regression Confidents (b), Wald Statistics, Significance (*p*), and Odds Ratios for Prediction of (A) Later Major Depressive Disorder (MDD) Diagnosis or (B) Later Anxiety Disorder Diagnosis

(A) Predicting Later MDD	b	Wald	*p* Value	Odds Ratio
Baseline irritability	*0.31*	*7.32*	*.01*	*1.37*
Sex (girls > boys)	−0.10	0.10	.75	0.91
Maternal MDD	0.63	4.08	.04	1.88
Baseline MDD symptoms	*0.54*	*37.76*	*<.001*	*1.71*
Adverse childhood experiences	*0.23*	*7.14*	*.01*	*1.26*
(B) Predicting Later Anxiety	**b**	**Wald**	***p* Value**	**Odds Ratio**
Baseline irritability	0.02	0.02	.89	1.02
Sex (girls > boys)	0.37	1.88	.17	1.45
Maternal anxiety	0.07	0.03	.86	1.07
Baseline anxiety symptoms	*0.44*	*38.61*	*<.001*	*1.55*
Adverse childhood experiences	0.03	0.13	.71	1.03

Note: Significant predictors (p < .05) are in italics. Baseline irritability was a significant predictor of later MDD above and beyond sex, maternal MDD, baseline MDD symptoms (subtracting out irritability symptoms from the depression module), and adverse childhood experiences. This effect was not significant when predicting anxiety.

TABLE S5 Associations Between Thickness and Covariates

	Left Superior Frontal Gyrus	Left Superior Temporal Gyrus	Right Inferior Parietal Lobule	Left Rostral Middle Frontal Gyrus	Right Superior Frontal Gyrus
Age at scan 1 (y)	−0.149	−0.143	−0.203*	−0.308**	−0.256**
Sex	−0.768	0.064	−0.758	0.059	0.26
Baseline non-irritability depression symptom severity	0.017	−0.021	0.009	−0.076	−0.099
Baseline income-to-needs ratio	−0.065	−0.058	0.058	0.009	−0.098
IQ	−0.100	−0.121	0.091	−0.051	−0.116
Baseline adverse life events	−0.044	−0.059	−0.106	−0.033	−0.094
Irritability severity at scan 1	0.116	0.049	0.030	−0.052	0.010
Maternal depression history	−0.295	−0.408	−0.46	0.337	0.353
Psychotropic medication use ever	−2.022*	−1.504	−0.221	−0.124	−1.512
ADHD diagnosis ever	−1.649	−0.403	0.837	1.714	0.657
MDD diagnosis ever	1.209	−0.372	0.327	1.652	1.18
Anxiety diagnosis ever	−0.441	0.565	0.674	1.097	0.462
ODD/CD diagnosis ever	−1.695	−1.101	0.093	0.981	0.309

Note: *This table displays associations between factors that varied by trajectory group or that could influence cortical thickness. T statistics are presented for effects of binary variables (sex, maternal depression history, medication use, and child diagnoses). Pearson correlation coefficients are presented for the remaining variables. ADHD = attention-deficit/hyperactivity disorder; CD = conduct disorder; MDD = major depressive disorder; ODD = oppositional defiant disorder.*
*$p < .05$; **$p < .01$; ***$p < 001$.*