



ARTICLE

Accelerated cortical thinning within structural brain networks is associated with irritability in youth

Robert J. Jirsaraie¹, Antonia N. Kaczurkin¹, Sage Rush¹, Kayla Piiwia¹, Azeez Adebimpe¹, Danielle S. Bassett^{1,2,3,4,5}, Josiane Bourque¹, Monica E. Calkins¹, Matthew Cieslak¹, Rastko Ciric¹, Philip A. Cook⁶, Diego Davila¹, Mark A. Elliott⁶, Ellen Leibenluft⁷, Kristin Murtha¹, David R. Roalf¹, Adon F. G. Rosen¹, Kosha Ruparel¹, Russell T. Shinohara⁸, Aristeidis Sotiras⁶, Daniel H. Wolf¹, Christos Davatzikos^{6,2,3} and Theodore D. Satterthwaite¹

Irritability is an important dimension of psychopathology that spans multiple clinical diagnostic categories, yet its relationship to patterns of brain development remains sparsely explored. Here, we examined how transdiagnostic symptoms of irritability relate to the development of structural brain networks. All participants ($n = 137$, 83 females) completed structural brain imaging with 3 Tesla MRI at two timepoints (mean age at follow-up: 21.1 years, mean inter-scan interval: 5.2 years). Irritability at follow-up was assessed using the Affective Reactivity Index, and cortical thickness was quantified using Advanced Normalization Tools software. Structural covariance networks were delineated using non-negative matrix factorization, a multivariate analysis technique. Both cross-sectional and longitudinal associations with irritability at follow-up were evaluated using generalized additive models with penalized splines. The False Discovery Rate ($q < 0.05$) was used to correct for multiple comparisons. Cross-sectional analysis of follow-up data revealed that 11 of the 24 covariance networks were associated with irritability, with higher levels of irritability being associated with thinner cortex. Longitudinal analyses further revealed that accelerated cortical thinning within nine networks was related to irritability at follow-up. Effects were particularly prominent in brain regions implicated in emotion regulation, including the orbitofrontal, lateral temporal, and medial temporal cortex. Collectively, these findings suggest that irritability is associated with widespread reductions in cortical thickness and accelerated cortical thinning, particularly within the frontal and temporal cortex. Aberrant structural maturation of regions important for emotional regulation may in part underlie symptoms of irritability.

Neuropsychopharmacology (2019) 44:2254–2262; <https://doi.org/10.1038/s41386-019-0508-3>

INTRODUCTION

Irritability is a debilitating dimension of psychopathology that cuts across multiple psychiatric disorders, including major depression, bipolar disorder, generalized anxiety disorder, oppositional defiant disorder, and disruptive mood dysregulation disorder. Symptoms of irritability include angry mood, exaggerated responses to negative stimuli, and limited frustration tolerance [1]. More specifically, irritability can be categorized as a chronically increased reactivity to negative emotional stimuli that is manifested verbally or behaviorally through temper tantrums, verbal rages, and/or aggression toward people or property. While many youths with severe irritability are diagnosed with bipolar disorder, studies utilizing careful assessment and longitudinal follow-up have demonstrated that the phenotype of chronic irritability is distinct from bipolar disorder [2–6]. This conceptualization is supported by evidence that symptoms of irritability exist along a continuum and can be measured reliably across disorders as a unique dimension [7–9]. Prior studies have revealed that

irritable youths are at increased risk for anxiety and depressive disorders and diminished functional outcomes, including reduced income and educational attainment [2, 10, 11]. Despite growing research on the impact of irritability, its neurobiological substrates remain sparsely investigated. This gap is critical, as understanding how aberrant patterns of brain development confer vulnerability to irritability is a prerequisite for targeted interventions that “bend the curve” of brain maturation to achieve better outcomes.

Symptoms of irritability are common in early childhood but generally decrease with age [12]. However, there are subgroups of children who exhibit chronically elevated levels of irritability throughout development [12–14]. To date, only two studies have investigated how irritability relates to longitudinal brain development [15, 16]. Adleman et al. (2012) reported that irritable youths had decreased gray matter volume in regions that support cognitive control and motor inhibition, such as the pre-supplementary motor area, insula, and dorsolateral prefrontal cortex. In contrast, Pagliaccio et al. (2018) reported that chronically

¹Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; ²Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA; ³Department of Electrical & Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁴Department of Physics & Astronomy, College of Arts & Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁵Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁶Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁷Section on Mood Dysregulation and Neuroscience, National Institute of Mental Health (NIMH), 9000 Rockville Pike, Bethesda, MD 20892, USA and ⁸Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA 19104, USA
Correspondence: Theodore D. Satterthwaite (sattertt@upenn.edu)

The preliminary findings from this study were presented at the 2018 Society of Biology Psychiatry Annual Meeting.

Received: 18 May 2019 Revised: 1 August 2019 Accepted: 14 August 2019

Published online: 2 September 2019

irritable youths had increased cortical thickness (CT) in the superior temporal and dorsolateral prefrontal cortex. Notably, longitudinal analyses from both studies did not detect any significant differences in the rate of cortical changes between the irritable group and typically developing groups. However, both studies used standard image analysis techniques and assessed irritability using a case-control design.

In addition to this prior work in structural imaging, studies using functional imaging have found that chronic irritability is associated with aberrant activation in fronto-temporal regions important for emotion regulation, including the dorsolateral prefrontal cortex, lateral orbitofrontal cortex, and anterior temporal cortex [17–19]. Additional functional magnetic resonance imaging (MRI) studies using response inhibition tasks have similarly reported that irritability is related to abnormal activation of the dorsolateral prefrontal cortex, which could further contribute to deficits in affect regulation [20, 21]. Studies using functional near infrared spectroscopy suggest that abnormalities of the dorsolateral prefrontal cortex can be detected as early as 3–5 years of age [22, 23]. Furthermore, anger and irritability were found to be inversely correlated with the functional connectivity between the amygdala and regulatory regions such as the orbitofrontal cortex [24–26].

Taken together, these findings suggest that abnormalities within executive and affective circuitry may be associated with chronic irritability. However, longitudinal studies of irritability are scarce, thereby limiting inferences that directly relate irritability to development. In fact, previous longitudinal studies on irritable youths have only acquired neuroimaging data between the ages of 7 to 16 years [15, 16]. Consequently, less is known about the neural markers of irritability during both infancy and early adulthood. Both stages of development are important given that the fronto-temporal regions critical for affect regulation undergoes protracted maturation throughout childhood and well into early adulthood [9, 27]. In particular, mapping neural trajectories as irritable youths transition into adulthood could be helpful for preventing deleterious outcomes during that period, which is when substance use, mood, and psychotic disorders often emerge [28].

In this study, we investigated how longitudinal structural brain maturation relates to a dimensional measure of irritability in a transdiagnostic sample of adolescents and young adults. We hypothesized that irritability at follow-up would be associated with reduced CT and accelerated longitudinal thinning in brain regions implicated in emotion regulation such as the dorsolateral, ventrolateral, and orbitofrontal cortex. In contrast to prior studies that examined structural differences within specific regions [15] or across hundreds of thousands of voxels [16], we studied structural covariance networks using a powerful multivariate analysis technique known as non-negative matrix factorization (NMF). Previous work using this approach has demonstrated that NMF-derived networks enhance both interpretability and statistical power [29, 30]. As described below, we provide new evidence that irritability is associated with reduced CT and accelerated thinning in brain regions critical for emotion regulation.

METHODS

Participants

A total of 137 participants (mean age = 16.0 years, $SD = 3.2$, range = 8–22 years, 80 non-Caucasian, 83 females) were recruited from the Philadelphia Neurodevelopment Cohort (PNC; [31, 32]). As part of the PNC, all participants received a psychiatric screening interview and a structural MRI scan. Four items from the psychiatric interview provided a baseline screening for irritability (see Supplementary Methods for details). Participants who endorsed symptoms of irritability from at least one of the four

binary GOASSESS items within the context of a psychiatric disorder were preferentially contacted about a follow-up study, in order to evaluate the dimensional impact of irritability across psychiatric disorders. Of the participants who were recruited for a follow-up visit, 102 endorsed at least one symptom of irritability and 35 did not screen positive for any symptoms at baseline (see Table S1). Follow-up diagnostic interviews, irritability assessments, and MRI scans were completed with a mean inter-scan interval of 5.2 years (inter-scan interval $SD = 1.0$, mean age = 21.1 years, $SD = 3.0$, range = 13–26 years). All procedures were approved by the Institutional Review Boards of the University of Pennsylvania and the Children's Hospital of Philadelphia.

Clinical phenotyping

Psychopathology symptoms at baseline were evaluated using a structured computerized screening interview [33], which was a modified version of the Kiddie-Schedule for Affective Disorders and Schizophrenia [34]. As previously described [35], the GOASSESS interview collected information on symptoms, frequency, duration, distress, and impairment for lifetime mood, anxiety, behavioral, eating, and psychosis-spectrum disorders. A total of 98 participants screened positive for at least one psychiatric disorder, and 39 did not screen positive for any disorder. Notably, a specialized assessment of irritability was not collected at baseline. However, four items from the GOASSESS screening interview were summed to provide a coarse dimensional assessment of irritability (see Supplementary Methods). All 98 participants with a psychiatric disorder reported symptoms of irritability at baseline along with 4 participants without a psychiatric disorder.

Clinical diagnoses at follow-up were assessed using a custom protocol [36] including modules from the Kiddie-Schedule for Affective Disorders and Schizophrenia [34], the Disruptive Mood Dysregulation Disorder Child Version [19], and the psychotic and mood differential diagnosis modules of the Structured Clinical Interview for DSM-IV [37]. All sections were administered by highly trained assessors in a semi-structured manner, allowing for follow-up probes and clarification of endorsed items. The semi-structured interview used to assess diagnosis status at follow-up was more detailed compared to the fully structured baseline screening interview (GOASSESS). As a result, the baseline interview may have been more sensitive but less specific, resulting in the higher rates of observed psychopathology at baseline. At the second time-point, there were 70 participants with at least one clinical diagnosis, and 67 participants who did not meet criteria for any disorders (see Table S2).

At follow-up, the Affective Reactivity Index (ARI) was used to measure self-reported symptoms of irritability [38]. This scale contains six items on symptom severity and a seventh item on functional impairment, which are rated on a three-point scale (ranging from 0 for “not true” to 2 for “certainly true”). The reliability and validity of the ARI have been previously reported [38, 39]. The sum of the first six items in the ARI was used as a dimensional measure of irritability, which was log transformed to obtain a more normal distribution (mean = 0.81, $SD = 0.75$, range = 0–2.7; see Figure S1A).

As psychiatric symptoms have substantial comorbidity, one potential risk is that any observed associations with irritability may be driven by collinearity with other symptom domains (see Fig. S1B). To evaluate this possibility, symptoms from relevant domains were measured at follow-up using the Beck Depression Inventory (BDI; [40]), the Screen for Child Anxiety Related Disorders (SCARED; [41]), the Strengths and Weaknesses of ADHD symptoms and Normal-behaviors (SWAN; [42]), and the PRIME Early Psychosis Screening Test (PRIME; [43]). These additional dimensions of psychopathology were included as covariates in sensitivity analyses.

Image acquisition

The baseline structural images were collected as part of the Philadelphia Neurodevelopment Cohort [31, 32]. In brief, all baseline scans were acquired on the same MRI scanner (Siemens TIM Trio 3 Tesla, Erlangen, Germany; 32-channel head coil) using the same imaging sequence. Structural brain imaging was completed using a magnetization-prepared, rapid acquisition gradient-echo (MPRAGE) T1-weighted image with the following parameters: TR 1810 ms; TE 3.51 ms; FOV 180 × 240 mm; matrix 256 × 192; 160 slices; slice thickness/gap 1/0 mm; TI 1100 ms; flip angle 9°; effective voxel resolution of 0.93 × 0.93 × 1.00 mm; total acquisition time 3:28 min. The follow-up structural scans were completed on a different MRI scanner (Siemens Prisma 3 Tesla, Erlangen, Germany; 64-channel head coil), but the imaging sequence and parameters were identical to those used on the baseline acquisition, and all measures were statistically harmonized prior to analysis; we describe this harmonization procedure in detail below.

Image processing and quality assurance

Structural image processing for estimating CT was completed with Advanced Normalization Tools version 2.1.0 (ANTs). A custom template and tissue priors were utilized to avoid registration bias and maximize sensitivity to detect regional effects that can be impacted by registration errors. Structural images were then processed and registered to this template using the ANTs CT pipeline [44, 45]. This procedure includes brain extraction, N4 bias field correction [46], Atropos probabilistic tissue segmentation [47], top-performing SyN diffeomorphic registration to template space [48], and diffeomorphic registration based CT measurement [49]. The ANTs CT pipeline has been validated in multiple open data sets, has good test–retest reproducibility, and enhanced ability to predict age and gender from regional thickness measurements when compared to FreeSurfer [44]. CT images were down-sampled to 2 mm voxels and subsequently smoothed by 4 mm before applying NMF.

All raw and processed images were reviewed by a highly trained imaging analyst who provided a quantitative score for each image [50]. This quality score was included as a covariate in all analyses, as described below. The analyst was trained to >85% concordance with faculty consensus ratings on an independent dataset. All images from both timepoints did not have any substantial artifacts. As an extra precaution to ensure that our observed results were not driven by motion artifact, we also re-computed analyses while controlling for motion during a resting-state fMRI scan acquired during the same scanning session (see Tables S3 and S4).

Non-negative matrix factorization

Cortical thickness data from both timepoints were summarized using structural covariance networks for two reasons. First, prior work has revealed inherent patterns of covariance in brain structure and analyzing the data according to this covariance structure enhances interpretability, as it has been shown to overlap with underlying functional networks [30, 50, 51]. Second, distilling the data into covariance networks reduces the large number of multiple comparisons present in voxel-based morphometry studies [51]. We used NMF to identify structural networks in which CT consistently co-varies across individuals and brain regions (see Fig. S2A). This tool (<https://github.com/asotiras/brainparts>) implements an extension of standard NMF that adopts orthonormality constraints for the estimated structural covariance networks and projective constraints for their respective loadings [52]. This procedure yields compact networks with positive weights that are more interpretable and reproducible compared to other decomposition techniques, such as principal component analysis and independent component analysis [29, 30]. Additionally, this procedure employs a

deterministic initialization algorithm [53] that leads to faster convergence, promotes sparsity, and ensures consistent results across multiple runs.

Multiple NMF resolutions were examined (2 to 30 networks, in increments of 2) in order to select the ideal number of components. Each resolution was computed in a split-half sample (matched on age, sex, image quality, and irritability) to evaluate reproducibility. The optimal number of components (24) was chosen based on the solution with the highest reproducibility (96.5%), which was quantified using the Adjusted Rand Index (see Fig. S2B). NMF networks are three-dimensional but were mapped to a surface for visualization using Caret (see Fig. 1; [54]).

Statistical harmonization

The CT values of structural brain networks from both timepoints were harmonized using the ComBat procedure to minimize the potential variance introduced by different MRI scanners. The ComBat procedure models scanner-specific scaling effects by utilizing Bayesian techniques that were originally implemented in the context of statistical genomics. These methods improve stability of the estimated parameters compared to traditional approaches, such as covarying for site/protocol effects [55]. Subsequent studies using neuroimaging data have shown that ComBat removes unwanted sources of scanner variability and preserves biological associations in the data [56–58]. The ComBat algorithm has even been effective for lifespan studies with broad age ranges, similar to our longitudinal design [55]. Effects of age, sex, and irritability were protected in the harmonization process.

Statistical analyses

As described in detail below, we evaluated cross-sectional relationships between irritability and CT at both timepoints, in addition to the longitudinal associations between irritability at follow-up and the annualized rates of change in CT. Statistical analyses were completed using R version 3.5.1 [59], and all code is publicly available (<https://github.com/PennBBL/jirsaraieStructuralIrritability>). Given that brain development is a non-linear process [28, 60], age was always modeled using penalized splines within a generalized additive model [61]. In this type of model, a penalty is assessed on nonlinearity using restricted maximum likelihood in order to avoid overfitting. The false discovery rate was controlled ($q < 0.05$) using the Benjamini–Hochberg procedure to correct for multiple comparisons.

As a first step, we examined the relationships at baseline between CT and the sum of the four irritability screening items. Second, we evaluated the relationships at follow-up between CT and irritability as measured by the ARI. Third, we evaluated how irritability at follow-up (using the ARI) related to longitudinal rates of change. Cross-sectional analyses included age, sex, and quality assurance ratings as covariates; longitudinal analyses included age at baseline, sex, and an average of the quality assurance ratings across timepoints as covariates. Longitudinal analyses evaluated the annualized rates of change in CT [62]. The total difference in CT between both timepoints was divided by baseline CT to get the total percent change ($\Delta\text{CT}/\text{CT}_{\text{TP1}}$), which was subsequently divided by the time between scans (ΔAge), yielding standardized the percent change of CT across all participants (CT_{Rate}):

$$\text{CT}_{\text{Rate}} = \frac{(\Delta\text{CT}/\text{CT}_{\text{TP1}})}{\Delta\text{Age}}$$

Sensitivity analyses were conducted on all significant results to ensure that effects were not being driven by maternal education, paternal education, race, diagnosis status, or other dimensions of psychopathology (see above). Each of these variables were individually included as covariates. Two additional analyses were

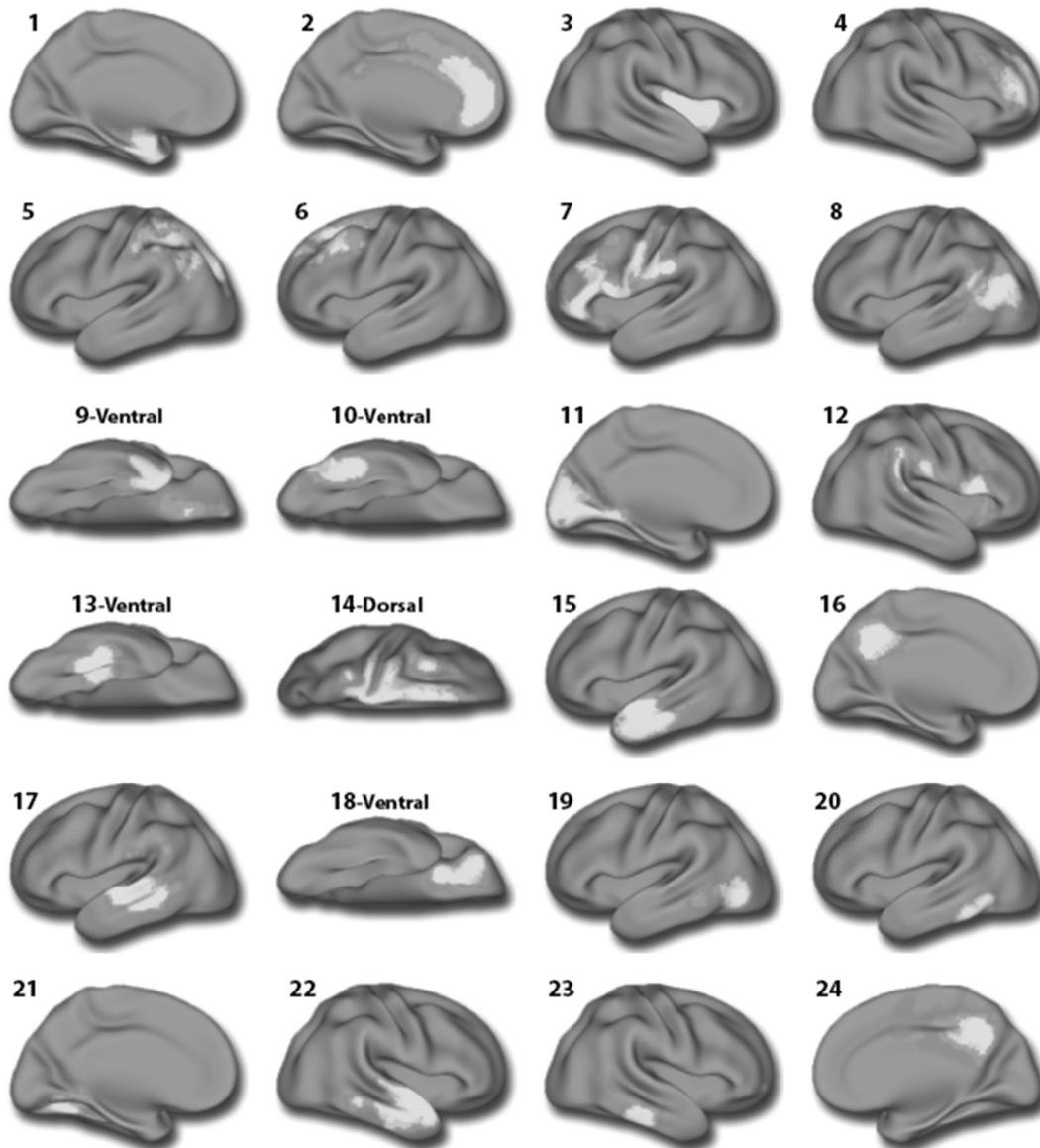


Fig. 1 Structural covariance networks. Visualization of the 24 structural covariance networks revealed by NMF. Warmer colors indicate higher loadings. Networks include (1) medial temporal lobe, (2) anterior cingulate, (3) insula, (4) anterior prefrontal cortex, (5) posterior parietal cortex, (6) dorsolateral prefrontal cortex, (7) inferior frontal cortex, (8) temporo-parietal junction, (9) temporal pole and medial orbitofrontal cortex, (10) inferior temporal gyrus, (11) occipital cortex, (12) peri-sylvian cortex, (13) fusiform gyrus, (14) dorsal frontoparietal cortex, (15) lateral temporal cortex, (16) posterior cingulate cortex, (17) supramarginal gyrus, (18) lateral orbitofrontal cortex, (19) temporo-occipital cortex, (20) inferior temporal cortex, (21) lingual gyrus, (22) superior temporal gyrus, (23) inferior temporal sulcus, (24) precuneus

conducted excluding participants taking psychotropic medications or using recreational drugs. Finally, we examined interactions between irritability and age, irritability and sex, as well as irritability and diagnostic status (typically developing vs. those with any psychiatric disorder). The models for cross-sectional, longitudinal, and age-related moderation analyses were as follows:

$$\begin{aligned} & \text{gam} \left(\begin{array}{l} \text{CT}_{\text{TP2}} \sim \text{Irritability}_{\text{TP2}} + \text{spline}(\text{Age}_{\text{TP2}}) + \text{Sex} \\ \quad + \text{Image Quality Rating}_{\text{TP2}} \end{array} \right) \\ & \text{gam} \left(\begin{array}{l} \text{CT}_{\text{Rate}} \sim \text{Irritability}_{\text{TP2}} + \text{spline}(\text{Age}_{\text{TP1}}) + \text{Sex} \\ \quad + \text{Image Quality Rating}_{\text{Mean}} \end{array} \right) \\ & \text{lm} \left(\begin{array}{l} \text{CT}_{\text{TP2}} \sim \text{Irritability}_{\text{TP2}} \times \text{Age}_{\text{TP2}} + \text{Sex} \\ \quad + \text{Image Quality Rating}_{\text{TP2}} \end{array} \right) \end{aligned}$$

RESULTS

Irritability at follow-up is elevated across diverse diagnostic categories

Irritability at follow-up was not associated with any demographic variables at follow-up, such as age ($p = 0.657$), sex ($p = 0.431$), race ($p = 0.092$), participant education ($p = 0.490$), paternal education ($p = 0.166$), or maternal education ($p = 0.872$). As expected, irritability at follow-up varied among diagnostic groups at both baseline ($p = 0.028$) and follow-up ($p < 0.001$). However, this effect was completely driven by differences between the typically developing participants and those who met diagnostic criteria for a psychiatric disorder; irritability did not differ between specific psychiatric disorders at baseline ($p = 0.535$) or follow-up ($p = 0.446$). Despite differing assessments of irritability between timepoints, the dimensional measures of irritability were significantly associated ($r = 0.211$, $p = 0.014$).

Cross-sectional irritability at follow-up is associated with thinner cortex
 As an initial step, we assessed cross-sectional associations between irritability and CT at each timepoint. At baseline, we examined the limited measure of irritability that was available (sum of four screening items), whereas the cross-sectional analysis in follow-up data used the specialized assessment provide by the ARI. Cross-sectional analyses of this coarse baseline data did not reveal any associations that survived FDR-correction. In contrast, dimensional irritability at follow-up was associated with thinner cortex within 11 of the 24 NMF networks (see Table 1). Effects were prominent in networks that included the orbitofrontal

cortex, which is critically involved in emotion regulation (see Fig. 2). Moderation analyses revealed that only the relationship between irritability at follow-up and the dorsolateral prefrontal cortex (network 6) weakened with age (see Figure S3); no other interactions were significant.

Sensitivity analyses provide convergent results
 Sensitivity analyses of cross-sectional findings at follow-up revealed that irritability remained associated with all 11 networks when controlling for maternal education, paternal education, diagnostic status, or race (see Table S5). Despite reduced statistical power, sensitivity analyses excluding 23 participants who were

Table 1. Cross-sectional associations between irritability at follow-up and structural covariance networks at follow-up

NMF-networks	<i>r</i>	β	SE	<i>t</i>	<i>p</i>	<i>p</i> -FDR
N-1: Medial temporal lobe	-0.29	-0.28	1.17	-4.21	<0.001	<0.001
N-8: Temporo-parietal junction	-0.23	-0.22	1.44	-3.01	0.003	0.009
N-9: Temporal pole & medial orbitofrontal cortex	-0.23	-0.20	1.30	-2.71	0.008	0.023
N-10: Inferior temporal gyrus	-0.20	-0.21	1.34	-2.76	0.007	0.018
N-11: Occipital cortex	-0.30	-0.26	1.71	-3.59	<0.001	0.003
N-13: Fusiform gyrus	-0.37	-0.35	1.07	-4.71	<0.001	<0.001
N-15: Lateral temporal cortex	-0.32	-0.31	1.10	-4.25	<0.001	<0.001
N-18: Lateral orbitofrontal cortex	-0.26	-0.26	0.84	-3.53	<0.001	0.004
N-20: Inferior temporal cortex	-0.20	-0.18	1.01	-2.54	0.012	0.027
N-21: Lingual gyrus	-0.20	-0.20	1.17	-2.56	0.012	0.026
N-22: Superior temporal gyrus	-0.22	-0.22	0.98	-3.02	0.003	0.009

Partial correlations (Pearson's *r*, controlling for age, sex, and data quality), β (standardized regression coefficients), standard errors, *t*-statistics, uncorrected and FDR-corrected *p*-values for significant associations between irritability and CT at follow-up. All analyses were conducted on a log transformation of irritability to improve normality and contained 132 degrees of freedom

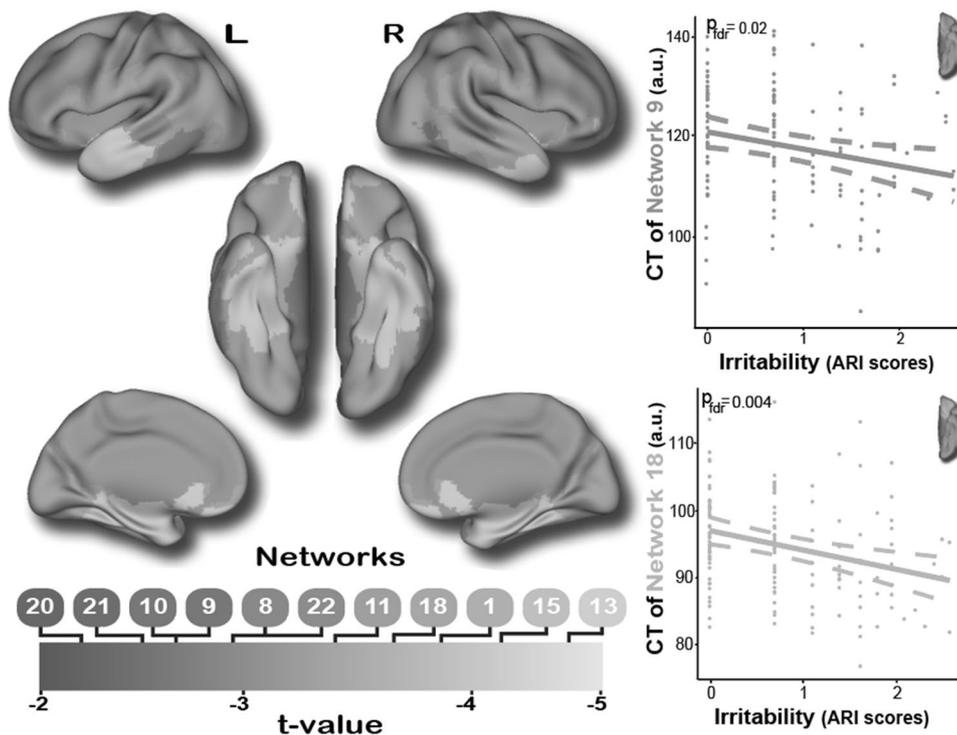


Fig. 2 Thinner cortex within multiple structural covariance networks is associated with irritability at follow-up. Cross-sectional analyses reveal that thinner cortex within multiple frontal and temporal networks were associated with irritability (log-transformed ARI). These networks comprise regions that are critical for affective regulation, such as medial temporal cortex, temporal pole, and both medial and lateral orbitofrontal cortex

Table 2. Longitudinal associations between irritability at follow-up and annualized rates of change in structural covariance networks

NMF-networks	<i>r</i>	β	SE	<i>t</i>	<i>p</i>	<i>p</i> -FDR
N-1: Medial temporal lobe	-0.35	-0.36	<0.01	-4.46	<0.001	<0.001
N-7: Inferior frontal cortex	-0.21	-0.21	<0.01	-2.48	0.014	0.043
N-9: Temporal pole and medial orbitofrontal cortex	-0.28	0.28	<0.01	-3.36	0.001	0.008
N-11: Occipital cortex	-0.35	-0.34	<0.01	4.32	<0.001	<0.001
N-15: Lateral temporal cortex	-0.22	-0.23	<0.01	-2.70	0.008	0.038
N-17: Supramarginal gyrus	-0.21	-0.21	<0.01	-2.53	0.013	0.043
N-18: Lateral orbitofrontal cortex	-0.20	-0.20	<0.01	-2.45	0.016	0.043
N-20: Inferior temporal cortex	-0.24	-0.23	<0.01	-2.81	0.006	0.034
N-22: Superior temporal gyrus	-0.20	-0.21	<0.01	-2.44	0.016	0.043

Partial correlations (Pearson's *r*; controlling for age, sex, and data quality), β (standardized regression coefficients), standard errors, *t*-statistics, uncorrected and FDR-corrected *p* values for significant associations between irritability at follow-up and annualized rate of change of CT. All analyses were conducted on a log transformation of irritability to improve normality and contained 132 degrees of freedom

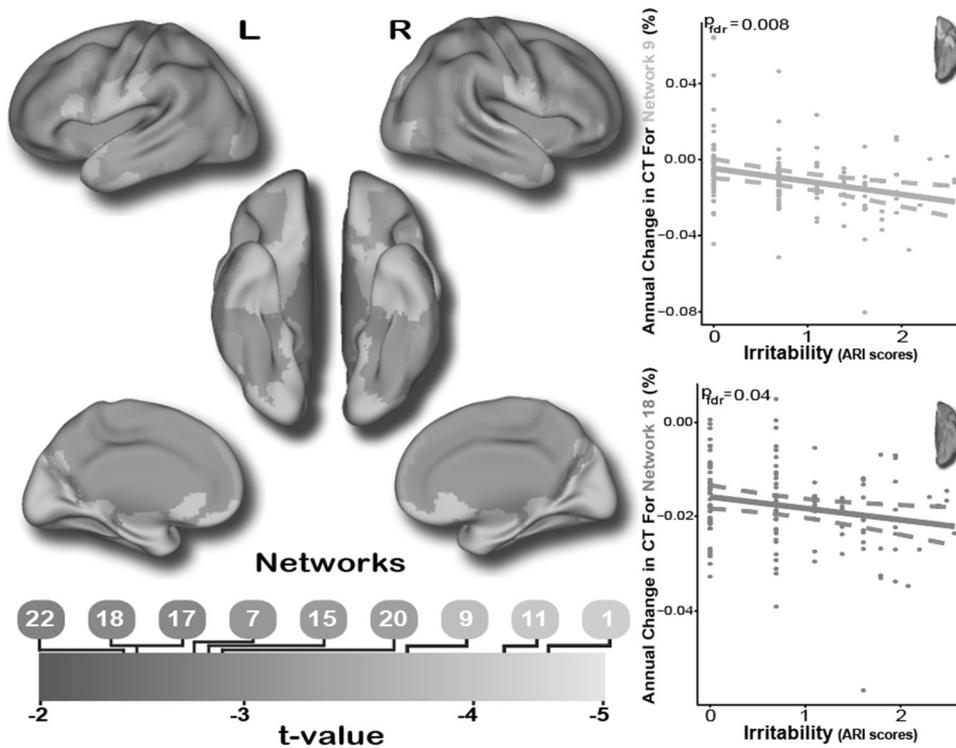


Fig. 3 Accelerated cortical thinning is associated with irritability. Longitudinal analyses revealed that the annual rates of change in many frontal and temporal networks were also associated with irritability (log-transformed ARI). Consistent with cross-sectional findings, these networks comprise the temporal pole, medial temporal cortex, and both medial and lateral orbitofrontal cortex

taking psychotropic medications did not significantly impact results, nor did excluding 32 participants who tested positive on a urine drug screen (primarily cannabis; see Table S5). Additional sensitivity analyses were conducted to determine whether these significant effects were specific to irritability or could reflect comorbid dimensions of psychopathology. Inclusion of depression, anxiety, or psychosis-spectrum symptoms as covariates did not significantly impact results; when dimensional ADHD symptoms were included as a covariate 5 of the original 11 networks persisted (see Table S5).

Irritability at follow-up is associated with accelerated longitudinal cortical thinning
 Finally, we evaluated associations between irritability at follow-up and the longitudinal rates of change within CT networks.

These analyses revealed that irritability at follow-up was significantly associated with accelerated longitudinal cortical thinning within nine networks (see Table 2 and Fig. 3). Notably, seven of these networks overlapped with the cross-sectional associations at follow-up. Sensitivity analyses revealed that irritability remained associated with rate of change within all nine CT networks when controlling for maternal education, paternal education, diagnostic status (except for network 18), or race (see Table S6). All significant effects remained when excluding participants who were taking psychotropic medications; four networks persisted after excluding participants who used had a positive urine drug screen (see Table S6). Finally, irritability remained associated with CT in these networks when controlling for comorbid symptoms of ADHD; four networks remained significant when covarying for anxiety or depression

(see Table S6). Age, sex, and diagnostic status did not moderate any longitudinal relationships.

DISCUSSION

We evaluated the dimensional association between irritability in youth and CT within structural covariance networks. Relative to previous studies on the neurodevelopment of irritability [15, 16], we observed widespread effects across the frontal and temporal cortex. Of the 24 covariance networks derived from NMF, irritability was associated with thinner cortex within 11 networks at follow-up and accelerated longitudinal thinning within 9 networks. In general, these relationships did not vary by age, sex, or diagnosis status, and some relationships could not be explained by confounding variables or comorbid dimensions of psychopathology. The most notable effects were observed in regions that are critical for emotion regulation, including orbitofrontal, lateral temporal, and medial temporal cortex. As described below, these results provide important insights regarding the neurodevelopmental substrates of transdiagnostic irritability in youth.

Previous functional MRI studies have found that irritability is associated with aberrations in multiple brain regions, including the orbitofrontal cortex, medial temporal cortex, temporal pole, superior temporal gyrus, and occipital cortex [17–19, 63]. We build upon this literature regarding brain function and demonstrate that many of these same regions exhibited CT reductions and accelerated cortical thinning that was related to transdiagnostic symptoms of irritability. Structural deficits in regions necessary for affect regulation, such as the orbitofrontal cortex and medial temporal cortex, may be critical for the pathogenesis of irritability. Additionally, abnormalities in the fusiform gyrus and visual cortex may relate to deficits in facial emotion recognition, which has been robustly associated with irritability in prior work [19, 64–67]. However, irritability was also associated with multiple other brain systems that were not predicted, including the motor and visual cortexes, which will need to be replicated by future studies. Taken together, these findings further support the notion that deficits in inter-related brain networks underlie vulnerability to irritable mood.

Previous structural and functional MRI studies have suggested that irritability is related to abnormalities within the dorsolateral prefrontal cortex [15–18, 22, 23]. Surprisingly, we did not find this association. However, previous studies were conducted on participants between the ages of 7 and 18, whereas our age range extended up to 26 years. Further, Tseng et al. (2018) reported that activation of the medial prefrontal cortex and anterior cingulate cortex during a frustration attention task was greater in irritable children compared to irritable young adults. Thus, it is possible that the relationship between irritability and prefrontal cortex does not extend into early adulthood. Moderation analyses modeling an irritability-by-age interaction supported this possibility: among the younger participants in our sample, irritability was negatively associated with CT of the dorsolateral prefrontal cortex. However, this relationship waned by early adulthood.

The current study documented more robust and widespread associations between irritability and CT than previous structural MRI studies [15, 16, 63, 68]. This difference may be attributed to three distinct advantages of the current approach. First, the structural covariance networks defined by NMF provided a parsimonious summary of the high-dimensional imaging data that limited multiple comparisons. This concise summary of the data allowed us to use rigorous FDR-correction for all comparisons, rather than cluster-based inferences that may be prone to Type 1 errors in many common implementations [69]. Second, instead of a categorical case-control approach, we used a

validated dimensional measure of irritability. While case-control studies are useful for describing DMDD as a distinct clinical syndrome [1], a dimensional measurement of irritability may provide a greater sensitivity to detect individual differences in brain structure [9]. Lastly, most studies have analyzed irritability within a specific psychiatric disorder, such as bipolar disorder [16] or depression [15]. In comparison, relatively fewer studies have examined irritability in the context of the comorbidity that is commonly observed in clinical practice [18, 19, 67, 68]. Our study included a transdiagnostic sample with substantial comorbidity, resulting in findings that may be more generalizable to the community.

Despite these advantages, several limitations should be noted. One disadvantage to having a sample with diverse clinical phenotypes and comorbid disorders is that we did not have sufficient statistical power to test whether the neural mechanisms underlying irritability differed among specific psychiatric disorders. Identifying the common and dissociable neural correlates of irritability across specific disorders could potentially have important implications for targeted therapies [1, 9]. Additionally, this study would have benefited from having the same dimensional measure of irritability at baseline and follow-up so that longitudinal analyses could have tested within-subject changes in irritability and CT. As previously noted, this is the first longitudinal study to examine brain maturation as irritable adolescents transition into adulthood. Ideally, longitudinal studies with three or more timepoints would map trajectories of brain development from childhood to adulthood.

In summary, we found that transdiagnostic symptoms of irritability were associated with widespread reductions in CT and accelerated cortical thinning in multiple brain networks. In particular, structural deficits were found in networks that support emotion regulation, including the orbitofrontal, medial temporal, and lateral temporal cortex. Future studies should incorporate repeated measurements of irritability and data from multi-modal imaging. Ultimately, such findings could allow for targeted interventions in youth at risk.

FUNDING AND DISCLOSURE

This work was supported by grants from the National Institute of Mental Health: R01MH107703, R01MH113550, K99MH117274, R01MH112847, R01EB025471, R01MH113565, R01MH11207, S10OD023495, and R01EB022573. Support for developing multi-variate pattern analysis software was provided by a seed grant by the Center for Biomedical Computing and Image Analysis (CBICA) at Penn. Additional support was provided by the Lifespan Brain Institute at the Children's Hospital of Philadelphia and Penn Medicine & the AE Foundation. The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41386-019-0508-3>).

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

REFERENCES

1. Leibenluft E. Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *Am J Psychiatry*. 2011;168:129–42.
2. Stringaris A, et al. Adult outcomes of youth irritability: a 20-year prospective community-based study. *Am J Psychiatry*. 2009;166:1048–54.
3. Stringaris A, et al. Youth meeting symptom and impairment criteria for mania-like episodes lasting less than four days: an epidemiological enquiry. *J Child Psychol Psychiatry*. 2010;51:31–8.

4. Biederman J, et al. Further evidence of unique developmental phenotypic correlates of pediatric bipolar disorder: findings from a large sample of clinically referred preadolescent children assessed over the last 7 years. *J Affect Disord*. 2004;82(Suppl 1):S45–58.
5. Stringaris A, et al. Pediatric bipolar disorder versus severe mood dysregulation: risk for manic episodes on follow-up. *J Am Acad Child Adolesc Psychiatry*. 2010;49:397–405.
6. Leibenluft E, et al. Chronic versus episodic irritability in youth: a community-based, longitudinal study of clinical and diagnostic associations. *J Child Adolesc Psychopharmacol*. 2006;16:456–66.
7. Leibenluft E, et al. Defining clinical phenotypes of juvenile mania. *Am J Psychiatry*. 2003;160:430–7.
8. Research Domain Criteria (RDoC). Toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;167:748–51.
9. Casey BJ, Oliveri ME, Insel T. A neurodevelopmental perspective on the Research Domain Criteria (RDoC) framework. *Biol Psychiatry*. 2014;76:350–3.
10. Brotman MA, et al. Prevalence, clinical correlates, and longitudinal course of severe mood dysregulation in children. *Biol Psychiatry*. 2006;60:991–7.
11. Dougherty LR, et al. Preschool irritability predicts child psychopathology, functional impairment, and service use at age nine. *J Child Psychol Psychiatry*. 2015;56:999–1007.
12. Belden AC, Thomson NR, Luby JL. Temper tantrums in healthy versus depressed and disruptive preschoolers: defining tantrum behaviors associated with clinical problems. *J Pediatrics*. 2008;152:117–22.
13. 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.
14. S. Wakschlag L, et al. Clinical implications of a dimensional approach: the normal: abnormal spectrum of early irritability. *J Am Acad Child Adolesc Psychiatry* 2015;54:626–34.
15. Pagliaccio D, et al. Irritability trajectories, cortical thickness, and clinical outcomes in a sample enriched for preschool depression. *J Am Acad Child Adolesc Psychiatry*. 2018;57:336.e6
16. Adleman NE, et al. Cross-sectional and longitudinal abnormalities in brain structure in children with severe mood dysregulation or bipolar disorder. *J Child Psychol Psychiatry*. 2012;53:1149–56.
17. Deveney CM, et al. Neural mechanisms of frustration in chronically irritable children. *Am J Psychiatry*. 2013;170:1186–94.
18. Tseng W-L, et al. Brain mechanisms of attention orienting following frustration: associations with irritability and age in youths. *Am J Psychiatry*. 2018;176:67–76.
19. Wiggins JL, et al. Neural correlates of irritability in disruptive mood dysregulation and bipolar disorders. *Am J Psychiatry*. 2016;173:722–30.
20. Deveney CM, et al. Neural recruitment during failed motor inhibition differentiates youths with bipolar disorder and severe mood dysregulation. *Biol Psychol*. 2012;89:148–55.
21. Singh MK, et al. Neural correlates of response inhibition in pediatric bipolar disorder. *J Child Adolesc Psychopharmacol*. 2010;20:15–24.
22. Li Y, et al. The neural substrates of cognitive flexibility are related to individual differences in preschool irritability: a fNIRS investigation. *Dev Cogn Neurosci*. 2017;25:138–44.
23. Perlman SB, et al. fNIRS evidence of prefrontal regulation of frustration in early childhood. *Neuroimage*. 2014;85(Pt 1):326–34.
24. Kringelbach ML. The human orbitofrontal cortex: linking reward to hedonic experience. *Nat Rev Neurosci*. 2005;6:691.
25. Goddard GV. Functions of the amygdala. *Psychol Bull*. 1964;62:89–109.
26. Fulwiler CE, King JA, Zhang N. Amygdala-orbitofrontal resting-state functional connectivity is associated with trait anger. *Neuroreport*. 2012;23:606–10.
27. Paus T, Keshavan M, Giedd JN. Why do many psychiatric disorders emerge during adolescence? *Nat Rev Neurosci*. 2008;9:947–57.
28. Giedd JN, et al. Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci*. 1999;2:861.
29. Sotiras A, Resnick SM, Davatzikos C. Finding imaging patterns of structural covariance via non-negative matrix factorization. *Neuroimage*. 2015;108:1–16.
30. Sotiras A, et al. Patterns of coordinated cortical remodeling during adolescence and their associations with functional specialization and evolutionary expansion. *Proc Natl Acad Sci USA*. 2017;114:3527–32.
31. Satterthwaite TD, et al. The Philadelphia Neurodevelopmental Cohort: a publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage*. 2016;124:1115–9.
32. Satterthwaite TD, et al. Neuroimaging of the Philadelphia neurodevelopmental cohort. *Neuroimage*. 2014;86:544–53.
33. Calkins ME, et al. The psychosis spectrum in a young U.S. community sample: findings from the Philadelphia Neurodevelopmental Cohort. *World Psychiatry*. 2014;13:296–305.
34. Kaufman J, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry*. 1997;36:980–8.
35. Calkins ME, et al. The Philadelphia Neurodevelopmental Cohort: constructing a deep phenotyping collaborative. *J Child Psychol Psychiatry*. 2015;56:1356–69.
36. Calkins ME, et al. Persistence of psychosis spectrum symptoms in the Philadelphia Neurodevelopmental Cohort: a prospective two-year follow-up. *World Psychiatry*. 2017;16:62–76.
37. First, M.B. and M. Gibbon, The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II), in *Comprehensive handbook of psychological assessment, Vol. 2: Personality assessment*. 2004, John Wiley & Sons Inc: Hoboken, NJ, US. p. 134–143.
38. Stringaris A, et al. The Affective Reactivity Index: a concise irritability scale for clinical and research settings. *J Child Psychol Psychiatry*. 2012;53:1109–17.
39. Mulraney MA, Melvin GA, Tonge BJ. Psychometric properties of the affective reactivity index in Australian adults and adolescents. *Psychol Assess*. 2014;26:148–55.
40. Beck AT, Steer RA, Brown GK. Beck depression inventory-II, San Antonio, Vol. 78; 1996. p. 490–8.
41. Birmaher B, et al. The screen for child anxiety related emotional disorders (SCARED): Scale construction and psychometric characteristics. *J Am Acad Child Adolesc Psychiatry*. 1997;36:545–53.
42. Swanson JM, et al. Categorical and dimensional definitions and evaluations of symptoms of ADHD: history of the SNAP and the SWAN Rating Scales. *Int J Educ Psychological Assess*. 2012;10:51. p
43. Miller TJ, et al. Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. *Schizophrenia Bull*. 2003;29:703–15.
44. Tustison NJ, et al. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *Neuroimage*. 2014;99:166–79.
45. Ciric R, et al. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *NeuroImage*. 2017;154:174–87.
46. Tustison NJ, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*. 2010;29:1310–20.
47. Avants BB, et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*. 2011;54:2033–44.
48. Klein A, et al. Evaluation of volume-based and surface-based brain image registration methods. *Neuroimage*. 2010;51:214–20.
49. Das SR, et al. Registration based cortical thickness measurement. *Neuroimage*. 2009;45:867–79.
50. Rosen AFG, et al. Quantitative assessment of structural image quality. *NeuroImage*. 2018;169:407–18.
51. Eklund A, Nichols TE, Knutsson H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci USA*. 2016;113:7900–5.
52. Yang Z, Oja E. Linear and nonlinear projective nonnegative matrix factorization. *IEEE Trans Neural Netw*. 2010;21:734–49.
53. Boutsidis C, Gallopoulos E. *Gallopoulos, E. Svd based initialization: a head start for nonnegative matrix factorization*. *Pattern Recognit*. 2008;41:1350–62. vol. 1350–1362:1350–62
54. Van Essen DC, et al. An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Inf Assoc*. 2001;8:443–59. p
55. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8:118–27.
56. Fortin J-P, et al. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*. 2018;167:104–20.
57. Fortin J-P, et al. Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*. 2017;161:149–70.
58. Yu M, et al. Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. *Hum Brain Mapp*. 2018;39:4213–27.
59. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2018. <http://www.R-project.org>.
60. Lenroot RK, et al. Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *Neuroimage*. 2007;36:1065–73.
61. Wood SN, Augustin NH. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecol Model*. 2002;157:157–77.
62. Cannon TD, et al. Progressive reduction in cortical thickness as psychosis develops: a multisite longitudinal neuroimaging study of youth at elevated clinical risk. *Biol psychiatry*. 2015;77:147–57.
63. Dennis EL, et al. Irritability and brain volume in adolescents: cross-sectional and longitudinal associations. *Soc Cogn Affect Neurosci* 2019.

64. Winston JS, O'Doherty J, Dolan RJ. Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage*. 2003;20:84–97.
65. Brotman MA, 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;167:61–9.
66. Guyer AE, et al. Specificity of facial expression labeling deficits in childhood psychopathology. *J Child Psychol Psychiatry*. 2007;48:863–71.
67. Kircanski K, et al. A latent variable approach to differentiating neural mechanisms of irritability and anxiety in youth. *JAMA Psychiatry*. 2018;75:631–9.
68. Gold AL, et al. Comparing brain morphometry across multiple childhood psychiatric disorders. *J Am Acad Child Adolesc Psychiatry*. 2016;55:1027–e3.
69. Mueller K, et al. Commentary: Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Front Hum Neurosci*. 2017;11:345–345.