

Large-Scale Functional Brain Networks of Maladaptive Childhood Aggression Identified by Connectome-Based Predictive Modeling

Karim Ibrahim (✉ karim.ibrahim@yale.edu)

Yale University School of Medicine <https://orcid.org/0000-0002-8205-6723>

Stephanie Noble

Yale University

George He

Yale University

Cheryl Lacadie

Michael Crowley

Yale University School of Medicine

Gregory McCarthy

Dustin Scheinost

Yale University <https://orcid.org/0000-0002-6301-1167>

Denis Sukhodolsky

Yale University School of Medicine

Article

Keywords:

Posted Date: March 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-356217/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Disruptions in frontoparietal networks supporting emotion regulation have been long implicated in maladaptive childhood aggression. However, the association of connectivity between large-scale functional networks in the human connectome with aggressive behavior has not been tested. By using a data-driven, machine learning approach, we show that the functional organization of the connectome during emotion processing predicts severity of aggression in children (n=129). Connectivity predictive of aggression was identified within and between large-scale networks implicated in cognitive control (frontoparietal), social functioning (default mode), and emotion processing (subcortical). Out-of-sample replication and generalization of findings predicting aggression from the functional connectome was conducted in an independent sample of children from the Adolescent Brain Cognitive Development study (n=1,791; n=1,701). These results define novel connectivity-based networks of child aggression that can serve as biomarkers to inform targeted treatments for aggression.

Introduction

Maladaptive aggression is among the most common reasons for referral to mental health services and spans across childhood psychiatric disorders, most notably disruptive behavior disorders, mood disorders, and autism spectrum disorder (ASD)¹. Neuroimaging studies have identified neural dysfunction in ventral and lateral prefrontal regulatory regions in children with aggressive behavior^{2, 3}. However, the search for brain-based predictors has yet to yield reliable neural biomarkers of child aggression that could inform treatment, such as by refining existing interventions, providing neural targets for novel treatments, and guiding assignment to interventions based on patterns of neural function⁴. Here, we use predictive modeling (machine learning) to examine patterns of brain-wide connectivity (i.e., functional connectivity across the entire brain or connectome that is based on region-to-region interactions) for predicting aggression in a transdiagnostic sample of children with aggressive behavior.

The emotion dysregulation model of aggression implicates disruptions in frontoparietal circuitry involved in the cognitive control of emotion⁵. Consistent with this view, meta-analytic studies^{3, 6} identified under-activity in the ventromedial and ventrolateral prefrontal cortices (vmPFC and vlPFC, respectively), dorsolateral prefrontal cortex (dlPFC), and temporal-parietal regions⁷⁻¹⁰. Neuroimaging studies of emotional processing also link childhood aggression with disruptions in functional connectivity between the amygdala and ventral prefrontal regions^{7, 10-12}. For instance, altered connectivity between the amygdala and ventral prefrontal regions was reported during emotion perception tasks in children with aggressive behavior^{7, 10, 11}. Recent functional MRI (fMRI) studies using emotional face perception tasks have also reported associations of irritability, an elevated propensity for anger and aggression, with reduced amygdala-vmPFC connectivity¹³ and reduced amygdala-dlPFC connectivity¹⁴ in children with aggressive behavior.

While a large body of neuroimaging work has elucidated key elements of the neural circuitry involved in aggression^{7,10,13,14}, it is poorly understood how functional interactions spanning multiple large-scale networks in the connectome are associated with aggressive behavior in children. Connectome modeling approaches, which leverage machine learning algorithms and include a cross-validation step^{15,16}, have been increasingly used to predict behavioral phenotypes from large-scale functional networks^{16–19}. Thus, these data-driven approaches may hold promise in identifying disruptions in the functional interactions among specific networks in the human connectome that are associated with aggressive behavior, potentially facilitating the discovery of brain-behavior associations and biomarkers. However, no study has examined whether the functional organization of the brain is predictive of aggression in children. The current study is the first to use a connectome predictive modeling approach to identify neural markers of childhood aggressive behavior.

Here, we employed a machine learning connectivity approach (also known as connectome-based predictive modeling, CPM)²⁰, to identify brain networks predictive of aggression. CPM leverages whole-brain functional connectivity data (“connectomes”)^{18,19} and enables cross-validation, which protects against overfitting^{4,21}. First, we applied CPM to fMRI data from an emotion processing task that was acquired in a well-characterized, transdiagnostic sample of 129 children with aggressive behavior and healthy control children without aggression. Connectome modeling was used to identify brain-behavior associations between a continuous measure of aggression and functional connectivity between and within large-scale networks in a transdiagnostic sample of children^{15,16}. Based on previous work on abnormal prefrontal connectivity in aggression^{3,6,7,11}, we hypothesized that abnormal connectivity within and between ventral and lateral prefrontal networks would be most predictive of individual differences in the severity of aggressive behavior. Then, we tested the replication and generalizability of the identified large-scale functional brain networks of maladaptive childhood aggression in an independent sample of children from the Adolescent Brain Cognitive Development (ABCD) study. The smaller, transdiagnostic sample (N = 129) served as the discovery dataset because it is enriched for the variable of interest (i.e., aggressive behavior). The larger ABCD cohort served as a replication dataset because it includes a heterogeneous sample of youths that is ideal for testing the replication and generalizability of brain biomarkers.

Results

Prediction of Aggression

Our hypotheses centered on large-scale networks in the functional connectome that predict severity of aggressive behavior in children. Therefore, we first trained and tested a predictive model for aggression in a transdiagnostic sample of children. The overall CPM model revealed that patterns of brain-wide connectivity predicted severity of aggressive behavior (combined positive and negative networks: $r=0.31$, RMSE = 9.05, $p=0.005$ via permutation testing) (**Fig. 1**). Follow-up comparisons controlling for potential covariates were conducted for age, IQ, sex, motion, and psychotropic medication

use, which demonstrated similar model performance ($r > 0.25$ and $p < 0.03$ for all models). The CPM model also remained significant after accounting for internalizing behaviors, social behavior impairments, and ADHD diagnosis ($r > 0.22$ and $p < 0.05$ for all models). Further details are provided in the **Supplemental Results**. Analyses were also repeated using 10-fold cross-validation and similar results were observed although, as expected with 10-fold versus leave-one-out cross-validation, the correlation coefficient was smaller ($r = 0.24$, $RMSE = 9.30$, $p = 0.01$).

Network Anatomy and Localization of Circuits

Consistent with previous CPM work^{18,19}, the predictive edges were distributed throughout the brain, particularly for positive networks, and included connections between frontal, parietal, occipital, and temporal lobes (**Fig. 2a**). The spatial extent of both positive and negative networks together included 396 edges (369 positive, 27 negative) or less than 1.1% of possible connections that predicted severity of aggressive behavior. Further inspection of these dense networks revealed that there were 126 ipsilateral connections in the right hemisphere, 98 ipsilateral connections in left hemisphere, and 172 connections between the right and left hemispheres. Networks predictive of aggression included significantly more long- than short-range connections (168 long-range; 95 short-range, $c^2 = 10$, $p = 0.002$). Highest-degree nodes (i.e., nodes with the most connections) for the positive network included: a bilateral dlPFC node with connections to limbic, temporal-parietal, sensorimotor, and other prefrontal nodes; a temporal node with connections to limbic, parietal, insula, subcortical, sensorimotor, and prefrontal nodes (**Fig. 2b**). Other high degree nodes for the positive network included nodes in the bilateral vmPFC, right vlPFC, and bilateral temporal poles. Highest-degree nodes for the negative network included a temporal-parietal node in the supramarginal gyrus with connections to limbic, parietal, and prefrontal nodes as well as with connections to cerebellar and subcortical nodes (**Fig. 2b**).

At the network level, **Fig. 2c** summarizes connectivity based on the number of connections within and between neural networks (e.g., frontoparietal, default mode, salience). For positive connectivity, within network connections were observed for several networks including the frontoparietal, default mode, medial prefrontal, and salience networks. Several large-scale, between network connections were observed: between frontoparietal and salience networks, which contributed the majority of edges to the positive network; between the medial frontal and default mode, sensorimotor, and salience networks; between the frontoparietal and default mode, salience, sensorimotor, and subcortical networks; and between default mode and salience, sensorimotor, and visual networks. The negative network included relatively more connections between the cerebellum and sensorimotor and salience networks. Overall between network connectivity was further characterized by within-network connections in frontoparietal networks and between network connections with the salience network (**Fig. 2c**).

Sensitivity Analyses

Virtual Lesioning Analysis. To test the hypothesis of ventral and lateral prefrontal involvement, we used a virtual lesioning analysis to evaluate the sensitivity of edges within and between the dlPFC versus

edges from every other node. However, to explore other regions that contributed to the model but were not included in our *a priori* hypotheses, we also examined the robustness of other high-degree nodes (i.e., most predictive) that were identified in the CPM model. Because these nodes were also predictive of aggression, we reasoned that this could be informative for developing biomarkers. Thus, we selected regions that emerged as high degree nodes, including the dlPFC, that occurred in all iterations of the connectome predictive model. First, we retained only high-degree nodes and all edges connected to these nodes, performing CPM using only these edges. The CPM predictive model remained significant even when restricted to each of the following high-degree nodes and related edges: bilateral dlPFC ($r=0.27$, $RMSE=9.0$, $p=0.002$), bilateral temporal pole ($r=0.22$, $RMSE=9.18$, $p=0.01$), right vlPFC ($r=0.21$, $RMSE=9.20$, $p=0.02$), bilateral parietal cortex ($r=0.23$, $RMSE=9.12$, $p=0.009$), bilateral occipital cortex ($r=0.24$, $RMSE=9.08$, $p=0.006$), and the right cerebellum ($r=0.25$, $RMSE=9.03$, $p=0.004$) (**Fig. 3**). Next, to check model robustness, we systematically removed each of the high-degree nodes (i.e., retaining all other edges). We found that the model predicting aggressive behavior remained significant across each of the high-degree nodes for the bilateral dlPFC ($r=0.21$, $RMSE=9.18$, $p=0.03$), bilateral temporal pole ($r=0.30$, $RMSE=9.06$, $p=0.0005$), right vlPFC ($r=0.30$, $RMSE=9.05$, $p=0.0004$), bilateral parietal cortex ($r=0.29$, $RMSE=9.12$, $p=0.0009$), bilateral occipital cortex ($r=0.29$, $RMSE=9.07$, $p=0.0006$), and the right cerebellum ($r=0.28$, $RMSE=9.12$, $p=0.0009$)—indicating the complexity of the functional connectome and contribution of nodes across the brain in the prediction of aggression. However, it is important to note that the predictive model retaining only the dlPFC nodes showed the greatest effect size while the predictive model lesioning the dlPFC nodes showed the smallest effect size, indicating the contribution of this region to predicting aggression.

Construct Specificity. We then tested the construct specificity of high-degree nodes in predicting aggression. Therefore, we conducted post-hoc tests retaining these high-degree nodes and all related edges (i.e., removing all other edges) to evaluate the robustness of nodes in predicting aggression in subgroups with high severity of co-occurring behaviors (internalizing symptoms, ADHD symptoms, and social impairments). A cut-off T score of >65 was used on standardized continuous measures (CBCL Internalizing Behaviors and CBCL Attention Problems scores, SRS-2 SCI total) to form subgroups because this represents the cut-off for a clinically significant level of psychopathology. The CPM model predicting aggression from dlPFC and its associated nodes remained significant with the dlPFC emerging as a high degree node despite co-occurring symptoms in subgroups with high severity of internalizing ($r=0.38$, $RMSE=6.46$, $p=0.005$), ADHD ($r=0.27$, $RMSE=6.54$, $p=0.03$), and social impairments ($r=0.26$, $RMSE=6.76$, $p=0.05$) (**Fig. 4a-c**). Other high-degree nodes did not demonstrate similar performance as the dlPFC across each of the co-occurring symptom subgroups (all other P s > 0.06) and are therefore less robust and not further discussed in the main article. However, for the interested reader, we present these findings for other high-degree nodes in **Table S3**.

To further assess construct specificity of the dlPFC nodes, we then tested whether the CPM model predicted internalizing and ADHD in the total sample ($N=129$). Even when dlPFC network connectivity predicted aggression (**Fig. 4**), it did not predict internalizing behaviors ($r=-0.015$, $RMSE=9.94$, $p=0.86$) or

ADHD ($r=-0.29$, $RMSE=5.42$, $p=1$), or social behavior impairments ($r=-0.48$, $RMSE=27.1$, $p=1$)—note that negative correlations indicate CPM model failure.

External Replication and Validation: Aggression Prediction

Out-of-Sample Replication. We then tested the replication of findings from the transdiagnostic sample indicating an association between the functional connectome and aggressive behavior in an independent sample of children from the ABCD study²². Therefore, we trained and tested a model in the ABCD sample to predict aggressive behavior using task-based fMRI from 1,701 children (920 females) using the stop signal task (SST) and from 1,791 children (958 females) using the emotional n-back task (EN-back) (age range 9-10 years). These tasks were selected for external replication and validation purposes because we reasoned that the SST and EN-back tasks tap frontoparietal and fronto-amygdala circuitry that are relevant to aggressive behavior²²⁻²⁴ and, similar to current study, the EN-back task stimuli included a set of happy, fearful, and neutral expressions drawn from the NimStim Stimulus Set²⁵. In addition, the SST taps the construct of response inhibition²³, which is implicated in disruptive behavior disorders in youths²⁶. For both tasks (EN-back, SST), the CPM model predicted aggressive behavior in the independent sample (EN-back: $r_s=0.10$, $RMSE=3.98$, $p < 0.001$; SST: $r_s=0.07$, $RMSE=4.01$, $p=0.002$) (**Fig. 5a**). Follow-up comparisons controlling for potential covariates were conducted for age, IQ, and sex demonstrated similar model performance for both the SST and EN-back tasks ($r > 0.1$ and $p < 0.01$ for all models). The CPM models also remained significant after accounting for ADHD and internalizing symptoms ($r > 0.7$ and $p < 0.01$ for all models). For completeness, model performance when including these covariates is provided in the **Supplemental Results**. Similar to the CPM model predicting aggression in the discovery transdiagnostic sample (**Fig. 1**), highest-degree nodes in ventral and lateral prefrontal regions as well as temporal-parietal regions emerged as highly predictive in the independent sample (EN-back task: bilateral dlPFC, right temporal pole, right frontal eye fields, right vmPFC; SST task: bilateral dlPFC, left temporal pole, bilateral supramarginal gyrus, right vmPFC) (**Fig. 5b-d**). At the network level, **Fig. 5e** summarizes connectivity within and between large-scale neural networks for the SST and EN-back tasks, which demonstrated similar patterns to the CPM model from the discovery sample. Overall, between-network connectivity was characterized by connections between medial frontal and frontoparietal, sensorimotor, default mode, and salience networks, which contributed the majority of edges to the positive network, as well as between the frontoparietal and subcortical, cerebellum, and salience networks.

Out-of-Sample Validation. We also tested the generalizability of findings by testing the ability of the identified networks in the transdiagnostic sample (**Fig. 1**) to predict aggressive behavior in an independent sample of children from the ABCD study and vice versa. First, we found that aggression network strength in the current, transdiagnostic sample predicted aggression severity in the independent sample (ABCD) for the EN-back ($r_s=0.06$, $p=0.007$), but not the SST ($r_s=-0.002$, $p=0.9$). Next, we tested the generalization of the aggression model developed in the independent dataset (ABCD study) to the current

study's discovery sample. We found that models developed using the EN-back and SST tasks both generalized to the current sample (EN-back: $r_s=0.18$, $p=0.03$; SST: $r_s=0.24$, $p=0.005$).

Discussion

This study is the first to apply connectome-based machine learning to identify brain-wide patterns of network connectivity underlying aggression in children. Two key findings from this study clarify associations between brain-wide functional connectivity and aggression in children. First, during implicit emotion processing of faces, brain-wide connectivity predicted severity of aggressive behavior. Specifically, dlPFC nodes emerged as highly predictive in the connectome models and were robust in sensitivity analyses. On a network level, connections within and between networks implicated in cognitive control (medial frontal, frontoparietal), social functioning (default mode, salience), and emotion generation/reactivity (subcortical, sensorimotor) emerged as significant predictors of aggression. Second, we further demonstrated that the same networks can be used to predict aggression in an independent, heterogeneous sample using similar tasks that tap into cognitive processes of response inhibition and working memory, often impaired in children with aggressive behavior.

The central role of the dlPFC in predicting aggression was consistent for models trained/tested in the discovery transdiagnostic sample of children and in the external validation dataset using the ABCD study. These findings are consistent with prior work reporting disruptions in dlPFC connectivity in children with aggression and irritability during tasks of emotion regulation^{14,27} and implicit emotion processing^{3,7,10,11}, spanning frontoparietal and fronto-amygdala circuitry. The dorsal and ventral prefrontal cortex connects to parietal and limbic regions, such as the amygdala, forming a frontoparietal and frontolimbic network, tightly coupled with cognitive control systems involved in emotion regulation and executive functioning²⁸. Therefore, projections between the dorsal and ventral prefrontal cortex and limbic regions are critical in dampening the acquisition and expression of negatively valenced emotions^{29,30}. Here, perturbed dlPFC connectivity during the processing of emotionally expressive faces could suggest disruptions in the modulation of emotion generation/reactivity systems in response to salient socioemotional stimuli associated with aggression in children.

It is important to note that the dlPFC nodes emerged as high-degree features—that is, the highest-ranking nodes that consistently contributed to models predicting aggression in all of the iterations—and remained robust predictors despite co-occurring symptoms of internalizing, ADHD and social impairments. While the dlPFC nodes alone predicted aggression severity in all of the follow up analyses, other nodes located in the salience, limbic, and sensorimotor networks are also noteworthy and showed predictive ability when the dlPFC nodes were left out. Given that the dlPFC was observed as the highest-ranking feature, it is possible that disruption to the dlPFC—a central hub of the emotion regulation circuitry with reciprocal connections to temporal-parietal and other cortical regions^{29,30}—may progress from local or modular disruptions before fanning out to other network hubs that are globally connected to frontoparietal networks involved in emotion regulation³¹. Thus, the involvement of other seemingly disparate networks

in predicting aggression could indicate compensatory effects (e.g., over-connectivity) in which other functionally connected nodes take over the role of perturbed nodes (i.e., dlPFC) to maintain an optimal level of overall function. This interpretation aligns with the ‘cascading network failure’ theory, which states that initial, local changes to the connectome are cascaded across the networks. Therefore, the diverse interconnectedness of the dlPFC may make this highly connected region a potential ‘vulnerability hot spot’ for disruptions in the functional connectome associated with aggression.

While our observations that lateral and ventral prefrontal circuitry emerged as highly predictive features in the CPM models was consistent with our a priori hypotheses, it is important to note that temporal-parietal regions across the brain also emerged as significant features including the temporal pole, sensorimotor regions, and supramarginal gyrus/inferior parietal cortex. In particular, it is noteworthy that the supramarginal gyrus/inferior parietal cortex emerged as a predictive feature of aggression. Parietal regions are implicated in memory and attention, and are recruited during cognitive control processes along with the ventral prefrontal cortex to modulate limbic reactivity^{30, 32, 33}. Functional and structural MRI studies have shown that perturbations in parietal networks are associated with disruptive behavior in youths^{2, 3, 34, 35}, which includes regions of the supramarginal gyrus^{34–37}. Because data-driven, connectomic approaches are advantageous for identifying highly predictive features, such as brain regions and networks, that were not hypothesized a priori, these approaches have potential to advance development of robust and generalizable biomarkers in childhood-onset psychiatric disorders^{15, 16}. Thus, our findings indicate the need to integrate additional circuits and systems that have not been emphasized in prior models of aggression, such as parietal networks, but can further understanding of the neural underpinnings of aggressive behavior in children. Together, the patterns of brain-wide associations observed here suggest that aggression psychopathology is supported by a widely distributed circuitry.

Several large-scale networks emerged in the CPM models of aggression including the salience and default mode networks—two neural circuits associated with attentional and social cognitive processes^{38–41} that share reciprocal functional connections⁴². Our findings are consistent with prior work implicating associations between disruptive behavior and aberrant connectivity and activity in salience^{10, 43, 44} and default mode^{44–47} networks in children. Reduced activation in social brain circuitry is associated with conduct problems in youths relative to controls during social cognitive tasks^{3, 48, 49}. Behavioral studies also show that impairments in social behavior, including emotion recognition and higher-order social cognitive processes such as theory of mind and empathy, are associated with childhood aggressive behavior^{50, 51}. Given the overlap of neural networks supporting social cognitive and emotion regulation processes⁵², disruptions in salience, default mode, and frontoparietal networks in children with aggression suggests impairment in the ability to detect salient social cues or inhibit inappropriate social responses, leading to aggressive responding. Thus, the salience and default mode networks and their component subsystems represent plausible substrates for deficits in emotion perception⁵³, empathy⁵⁴, and moral reasoning⁵⁵ that have been observed in children with disruptive behavior disorders.

Multiple large-scale networks were involved in predicting aggression, emphasizing the complexity of the functional connectome in predicting behavioral phenotypes as discussed previously^{18, 19, 56}. Figure 6 shows a theoretical network model of maladaptive aggression, integrating the network-level connections we observed. Based on these findings, we propose that aggression is predicted by: 1) disruptions in coordination between a cognitive control network system implicated in emotion regulation and executive functioning involving hyper-connectivity between frontoparietal and medial prefrontal networks; 2) disruptions in the integration of an emotion generation/reactivity network involving hyper-connectivity between subcortical and sensorimotor networks; and 3) disruptions in coordination between a social functioning system involving hyper-connectivity between the salience and default mode networks, and integration of a social functioning system with cognitive control and emotion generation networks. The current model builds on previous triadic model accounts of aggression, suggesting disruptions in limbic circuitry in response to emotional stimuli coupled with disruptions in prefrontal regulatory circuitry, and altered connectivity between these systems²⁶, but also incorporates sensorimotor and salience networks that could provide a theoretical framework for future research.

This connectome-based model of aggression was also found to be replicable and generalizable to an out-of-sample dataset. First, replication of findings in ABCD indicated a similar pattern of brain organization across large-scale networks that predicted aggression severity. Second, the model from the in-house, discovery dataset predicted aggression in the external ABCD dataset and vice versa. We also observed a similar pattern of results implicating frontal, temporal-parietal, and sensorimotor networks in aggression. In particular, dlPFC nodes were highly predictive for both EN-back and SST tasks. It is important to note that the connectome model was replicated and generalized to the independent ABCD sample, despite different measures of aggression. Here, the RPQ was used as the primary measure of maladaptive aggression because children were selected based on a criterion cut-off using the CBCL Aggressive Behavior scale as part of the RDoC framework to identify aggression transdiagnostically. In the ABCD sample, the CBCL was the most optimal measure of aggression and assessed a similar construct as the RPQ. Given that the connectome model demonstrated replication and generalization of findings to ABCD, this demonstrates potential robustness of brain-based models to predict aggression using different but related behavioral measures of aggression. This suggests that transdiagnostic, generalizable neural systems could be implicated in childhood maladaptive aggression.

This study has several strengths, including the use of a novel brain-wide predictive modeling approach, out-of-sample replication and generalization, and a well-characterized sample of children with aggressive behavior. Results were also independent of confounds including age, IQ, motion, internalizing and ADHD behaviors, and ASD-related social behavior impairments. Nonetheless, some limitations should be noted. First, the present study is cross-sectional and future research is needed to understand how these findings may relate to trajectories of connectivity impairment implicated in aggression. Second, while models were robust to common co-occurring internalizing, ADHD, and ASD symptoms, the functional significance of the identified networks in relation to other forms of developmental psychopathology associated with aggression remains to be tested in larger samples. Third, our models have modest effect sizes, capturing

approximately 10% of the variance, which was reduced to 1–5% in the replication analysis. The small effect sizes observed here for replication and external validation could in part be due to the use of data-driven, machine learning approaches, which are expected to have small effect size estimates when applied to large datasets⁵⁷. Additionally, the effect sizes observed here are consistent with recent work examining brain-wide associations with phenotypic data in the ABCD dataset⁵⁸. Thus, it is possible that imaging-derived metrics in predicting modest amounts of variance in a complex phenotype may become a common finding with machine learning approaches that leverage brain-wide connectivity in large-scale samples. This could be partly due to the heterogeneity of complex neurocognitive traits and psychiatric disorders associated with brain-wide associations. Alternatively, given that psychopathology is supported by widely distributed circuitry⁵⁹, machine learning approaches applied to brain-wide associations may pick up on generalizable features inherent in the connectome that can predict a complex phenotype, potentially reflecting more conservative effect size estimates^{57, 60}. Future studies examining connectome predictive modeling approaches in large-scale samples will be important to assess the clinical utility of this method.

Conclusion

This study demonstrates that patterns of brain-wide connectivity can predict severity of aggressive behavior in children, even when controlling for co-occurring psychiatric symptoms of ADHD, internalizing and social behavior deficits. Networks predictive of aggression included medial frontal, default mode, frontoparietal, salience, sensorimotor, and subcortical networks as well as complex patterns of integration among these networks. The predictive ability of these networks also translated to an independent sample of children. These data demonstrate that individual differences in connectivity within large-scale functional networks implicated in cognitive control, social processing, and emotion generation/reactivity processes contribute to variability in aggression. These networks or “neural fingerprints” may be an appropriate target for development of biomarkers that can inform novel, targeted interventions for children with aggressive behavior.

Methods

Participants

This study included 100 children with aggressive behavior (27 females) and 29 typically developing healthy controls (HC group; 13 females). All participants were aged 8 to 16 years. Sample characteristics are shown in **Table 1**. Children with aggressive behavior were recruited for a treatment study of cognitive-behavioral therapy for irritability/anger and aggression⁶¹. Cross-sectional, functional MRI data was collected at baseline, prior to initiating the treatment. One of the inclusion criteria for the treatment study was a T-score of 65 or greater on the Aggressive Behavior Scale of the Child Behavior Checklist (CBCL)⁶², which is 1.5 standard deviations above the mean in the standardization sample and represents a cut-off for a clinically significant level of aggression. The Aggressive Behavior scale includes

16 items reflecting anger outbursts, verbal and physical aggression. HC participants were required to have no current or past history of psychiatric or neurological disorders and a CBCL Aggressive Behavior T-score below 55. Additional details regarding inclusion and exclusion criteria are provided in the **Supplement**. Consistent with an RDoC approach to investigating dimensions of psychopathology transdiagnostically, the sample included children with a range of psychiatric diagnoses (see **Table 1**). However, untreated PTSD and severe depression were exclusionary criteria because these disorders may present with pressing treatment needs, not targeted in the treatment study. Primary diagnosis reflected the chief symptoms for which youths sought treatment. All children also met criteria for Oppositional Defiant Disorder (ODD), Conduct Disorder, or Disruptive Mood Dysregulation Disorder (DMDD) based on a semi-structured diagnostic interview⁶³. The sample included 18 subjects with ASD and high levels of aggression; ASD diagnosis was confirmed by the *Autism Diagnostic Interview-Revised*⁶⁴ and *Autism Diagnostic Observation Schedule-2nd edition*⁶⁵. Children were recruited from the outpatient child psychiatry clinic at the Yale University Child Study Center and from outreach to the local schools, pediatricians and mental health providers. Healthy control children were recruited from the community via advertisements. Each participant's parent provided informed consent according to specifications by the institutional review board at the Yale University School of Medicine. Each child provided verbal and written assent.

Measures

The Reactive-Proactive Aggression Questionnaire (RPQ) total score⁶⁶ was used as the primary continuous measure of severity of aggressive behavior in CPM analyses. The RPQ is a 23-item parent-report scale that measures proactive and reactive aggression on a 3-point Likert scale. The RPQ Aggression Total score was used because it reflects the overall severity of maladaptive aggressive behavior in pediatric populations^{1, 67-69}. Children received a comprehensive diagnostic evaluation that included the *Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL)*⁶³ that was conducted with the parent and child by an expert clinician to establish DSM-5 diagnoses of Disruptive Behavior Disorders and co-occurring psychopathology. Full scale IQ was evaluated with the *Wechsler Abbreviated Scale of Intelligence (WASI)*⁷⁰. The CBCL Internalizing Behavior Problem score was used as a continuous measure of internalizing symptoms. The parent-rated Social Responsiveness Scale-Second Edition (SRS-2) Social Communication and Interaction subscale score (SCI; 53 items)⁷¹ was used as a dimensional measure of severity of social impairments. The SRS-2 SCI subscale represents four dimensions of social behaviors including social awareness, motivation, communication, and cognition. The SRS was initially developed as a measure of social impairment in ASD, but was shown to capture social cognitive impairments in the general population as well as in children with aggressive behavior⁷². Thus, we reasoned that the SRS-2 SCI subscale would provide an approximation of impairments in social functioning related to maladaptive aggression in children⁷³⁻⁷⁵. Additional detail related to study measures is provided in the **Supplement**.

Task

Participants completed a previously-validated fMRI task of emotionally expressive faces. Additional task details are also reported elsewhere⁶¹. The task uses a pseudorandomized block design with 12 blocks that each contain two randomly selected faces exhibiting the same expression: 6 calm emotion and 6 fearful emotion blocks. The face-expression pair images are randomly selected throughout the blocks and no individual face-expression image is displayed more than once throughout the paradigm (**Fig. S1**). Each block was 12 seconds in length and consisted of two faces displayed for 5.5 seconds each that were separated by a 1 second intertrial fixation cross. Blocks were separated by a jittered interblock interval between 8 and 12 seconds to optimize statistical efficiency. The interblock intervals were pseudorandomized such that the mean of all interblock intervals was 10 seconds. The first block was preceded by a 10 second initial fixation cross and the final block was succeeded by an identical 10 second fixation cross. The total duration of the paradigm was 284 seconds. Participants were instructed to perform a gender identification task using a button press in their left or right hand to indicate male or female, respectively. We examined connectivity across both emotions (fearful and calm) to understand broader deficits in emotion face processing associated with aggressive behavior, and to ensure sufficient continuous voxel timecourse for CPM analyses. Prior to the fMRI session, a mock scanner was used to acclimate participants to the scanning environment (see **Supplemental Methods: Training in the Mock Scanner**).

We conducted a series of repeated-measures analysis of covariance (ANCOVA) to examine the associations between aggression and task behavioral measures of accuracy and reaction time. Results for behavioral performance on the emotion face processing task are shown in the **Supplemental Results and Table S2**.

Data Acquisition and Preprocessing

Functional MRI was performed using a Siemens MAGNETOM Tim Trio 3 Tesla scanner with an upgrade for echoplanar images (EPI). A T1-weighted high-resolution anatomical scan was obtained for each participant for co-registration purposes: repetition time (TR)=2530 ms; echo time (TE)=3.31 ms; 1mm isotropic voxels; 176 slices; flip angle=7°; matrix size=256²; field of view (FOV)=256 mm. For each participant, 137 interleaved, oblique whole-brain functional volumes were collected in the axial plane parallel to the intercommissural (AC-PC) line using an EPI gradient echo sequence: TR=2000 ms; TE=25 ms; voxel size = 3×3×4 mm, 34 slices; flip angle=60°; matrix size=64²; FOV=220² mm.

Standard preprocessing and functional imaging statistical analyses were conducted using the Functional MRI of the Brain (FMRIB) Software Library (FSL Version 4.1.6; FMRIB, Oxford, United Kingdom)^{76, 77}. The first four volumes of each functional run were discarded to allow for the magnetization to reach a steady state. Functional data were temporally realigned to correct for interleaved slice acquisition. Motion was corrected using FSL MCFLIRT linear realignment tool⁷⁸. Of 106 potential scans in the aggressive behavior group and 30 in the healthy controls group, 6 children with aggressive behavior and one typically developing child were excluded from the final analysis owing to excessive motion and computer error during the scan, respectively. Thus, 129 scans were included in the

final analysis (100 in the aggressive behavior group and 29 in the healthy control group). Data were spatially smoothed with a 5-mm full width at half maximum isotropic Gaussian kernel with a non-linear high-pass filter (60 s cut-off). Individual participant analyses were conducted using FSL FMRI Expert Analysis Tool (FEAT). Functional images were registered to co-planar images, which were then registered to the high-resolution T1-weighted images and normalized to the Montreal Neurological Institute 152 template. In addition, we controlled for frame-to-frame displacement in our CPM analysis using partial correlation. For each participant, motion corrected data were transformed to standardized MNI space. Several covariates of no interest were regressed out from the data including the 12 motion parameters (six rigid body motion parameters and six temporal derivatives), mean white matter signal, mean cerebrospinal fluid signal, mean global signal, and the linear, quadratic, and cubic drifts. No significant between-group differences were observed in mean head motion detected during the functional scan (all $P_s > 0.37$) (**Table S1**).

Connectivity Matrices

Whole-brain functional connectivity analyses were conducted using BiImage Suite and previously described methods^{16, 79}. Network nodes were defined using the Shen 268-node brain atlas, which includes the cortex, subcortex, and cerebellum as described in prior work^{20, 56}. The atlas was warped from MNI space into single-subject space. Task connectivity was calculated on the basis of the task time courses^{18, 19}. This involved computation of the mean time courses for each of the 268 nodes (i.e., averaging the time courses of all constituent voxels). Node-by-node pairwise correlations were computed, and Pearson correlation coefficients were Fisher z-transformed to yield symmetric 268 × 268 connectivity matrices, in which each element of the matrix represents the connectivity strength between two individual nodes (i.e., “edge”).

Connectome-Based Predictive Modeling (CPM)

CPM was conducted to predict RPQ scores using published MATLAB scripts²⁰. A schematic diagram is shown in **Fig. S2**. CPM uses connectivity matrices and behavioral data (here, aggressive behavior severity) from individuals as input to generate a predictive model of the behavioral data from connectivity matrices. Edges and behavioral data from the training data set are correlated using regression analyses (here, Pearson’s correlation or—when controlling for other behavioral variables or covariates—partial correlation) to identify positive and negative predictive networks. Positive networks are networks for which increased edge weights (increased connectivity) are associated with the variable of interest, and negative networks are those for which decreased edge weights (decreased connectivity) are associated with the variable of interest. Single-subject summary statistics are then created as the sum of the significant edge weights in each network, entered into predictive models assuming linear relationships with behavioral data. The identified predictive networks and summary score model from the training data are then applied to the test data set to predict behavior. In the case of leave-one-out cross-validation used here, a single participant’s predicted value is generated by taking the data from all other participants as the training data set in an iterative manner until all participants have a predicted value.

Localization of predictive networks

Predictive networks identified using CPM are complex and composed of multiple brain regions and networks. Similar to previous CPM studies, predictive networks were summarized at multiple levels of data reduction including at the edge, node, network level¹⁶. Macroscale brain regions (e.g., prefrontal cortex, cerebellum) were based on anatomical labels presented in Finn et al.¹⁷. Canonical functional network localizations (e.g., frontoparietal, sensorimotor) were based on the functional networks presented in Nobel et al.⁸⁰ Additionally, for each node, the network theory measure degree was calculated as the sum of the number of edges for each node that belonged to the predictive networks. Finally, networks were summarized based on length-of-connection (short- versus long-range connectivity). Euclidean distance between the centroids of each brain region in the Shen atlas was used to classify short and long-range edges. First, distance was calculated

for each pair of regions as: $\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}$, where (x_1, y_1, z_1) and (x_0, y_0, z_0) represent

the centroid for any two regions. Pair-wise distances were median-separated into short- and long-range connections. Visualizations of predictive edges were created using BioImage Suite Web (<https://bioimagesuiteweb.github.io/webapp/connviewer.html>).

Significance of CPM Performance

For connectome analyses, the correspondence between predicted and actual values, or model performance, was assessed using Pearson's correlation (r) and root mean square error (defined as:

$$RMSE(\text{predicted}, \text{actual}) = \sqrt{(1/n \sum_{i=1}^n (\text{actual}_i - \text{predicted}_i)^2)}.$$

When using cross-validation, analyses in the left-out folds are not independent, and the number of degrees of freedom is thus overestimated for parametric p values. Therefore, permutation testing was performed. To generate null distributions for significance testing, we randomly shuffled the correspondence between behavioral variables and connectivity matrices by permuting subject assignments for behavioral variables 1,000 times and re-ran the CPM analysis with the shuffled data. Based on these null distributions, the p -values for predictions were calculated¹⁸. As we hypothesize a positive association between predicted and actual values, one-tailed p -values are reported.

Follow-up analyses were conducted to test the robustness of nodes that emerged as highly predictive in the CPM model to confounds (see **Supplemental Methods**). First, a virtual lesion analysis was used to test model performance after 'lesioning' or restricting CPM to high-degree nodes and their corresponding edges identified in the main model. Next, to test construct specificity of the dlPFC and other high-degree nodes in predicting aggression, we conducted post-hoc tests retaining the high-degree nodes and all edges connected to these nodes (i.e., removing all other edges), to evaluate the robustness of these

networks in predicting aggression in subgroups with high severity of co-occurring behaviors (internalizing symptoms, ADHD symptoms, and social impairments). A cut-off T score of >65 was used on standardized continuous measures (CBCL Internalizing Behaviors and CBCL Attention Problems scores, SRS-2 SCI total) to form subgroups because this represents the cut-off for a clinically significant level of psychopathology.

External Replication and Validation

We tested the replication of findings and generalizability of the aggregate model from the transdiagnostic discovery dataset (i.e., edges present in >95% of cross-validation folds) in predicting aggressive behavior in an independent, open-access sample of youths from the Adolescent Brain Cognitive Development (ABCD) study²². We applied the CPM aggression network models to an independent sample from the ABCD study of task-based fMRI from 1,701 children (920 females) using the stop signal task (SST) and from 1,791 children (958 females) using the emotional n-back task (EN-back) (age range 9-10 years). We selected these tasks for replication and validation purposes for the following reasons. First, the SST and EN-back tasks tap frontoparietal and fronto-amygdala circuitry that are relevant to aggressive behavior²²⁻²⁴. Second, similar to our study, the EN-back task stimuli included a set of happy, fearful, and neutral expressions drawn in part from the NimStim Stimulus Set²⁵. Additionally, the SST taps the construct of response inhibition²³ that is also implicated in disruptive behavior disorders in children²⁶. ABCD data was processed in a similar manner to the main dataset and is described in our previous work⁸¹. CPM was conducted as above with leave-one-out cross-validation. We computed Spearman's correlations (r_s) between predicted and observed CBCL aggressive behavior scores based on connectivity matrices. Spearman's correlations were used for ABCD to address the comparability in connectome prediction using different but related aggression measures between the discovery and replication/external datasets. We first tested the replication of findings and whether the functional organization of the connectome predicts aggressive behavior in ABCD. We then tested the generalizability of predictive models from the transdiagnostic, discovery dataset (N=129) to ABCD and vice versa.

Data Availability

To promote data transparency, anonymized data that support the findings of this study are available from the corresponding authors upon reasonable request. Data from the studies reported in this paper (transdiagnostic sample from Yale and ABCD) have also been shared on the National Institute of Mental Health Data Archive (NDA) <https://nda.nih.gov/>.

Declarations

Ethics: This study was reviewed and approved by the local ethical committee (institutional review board at the Yale University School of Medicine), and it was conducted in accordance with the declaration of Helsinki.

Acknowledgments

This work was supported by NIMH grant R01MH101514 (D.G.S.) and NICHD grant R01HD083881 (D.G.S.). K.I. is a fellow on NCATS grant TL1 TR001864 and the Translational Developmental Neuroscience Training Program (T32 MH18268) directed by Dr. Michael Crowley. We thank Ms. Sonia Rowley for her assistance with reviewing the final version of the manuscript, Ms. Iciar Iturmendi for her assistance with formatting figure 3, Dr. Megan Tudor for subject characterization assessments, and Ms. Emilie Bertschinger, Ms. Tess Gladstone and Ms. Carolyn Marsh for study coordination.

ABCD acknowledgement

A portion of the data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (<https://abcdstudy.org>), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093, and U01DA041025. A full list of supporters is available at <https://abcdstudy.org/federal-partners.html>. A listing of participating sites and a complete listing of the study investigators can be found at <https://abcdstudy.org/scientists/workgroups/>. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

Authorship Contributions

Conception: D.G.S., K.I., D.S., M.J.C., G.M.; Design of the work: D.G.S., K.I., D.S., G.M.; Acquisition, analysis, interpretation of data: D.G.S., K.I., D.S., S.N., C.L., G.H.; Manuscript writing and revising: D.G.S., K.I., D.S., M.J.C., G.M., S.N., C.L., G.H.

Competing Interests

The authors declare no competing interests.

References

1. Connor, D.F., *et al.* Maladaptive aggression: with a focus on impulsive aggression in children and adolescents. *J. Child Adolesc. Psychopharmacol.* **29**, 576–591 (2019).
2. Rogers, J.C. & De Brito, S.A. Cortical and subcortical gray matter volume in youths with conduct problems: a meta-analysis. *JAMA psychiatry* **73**, 64–72 (2016).

3. Alegria, A.A., Radua, J. & Rubia, K. Meta-analysis of fMRI studies of disruptive behavior disorders. *Am. J. Psychiatry* **173**, 1119–1130 (2016).
4. Gabrieli, J.D., Ghosh, S.S. & Whitfield-Gabrieli, S. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron* **85**, 11–26 (2015).
5. Davidson, R.J., Putnam, K.M. & Larson, C.L. Dysfunction in the neural circuitry of emotion regulation—a possible prelude to violence. *Science* **289**, 591–594 (2000).
6. Noordermeer, S.D., Luman, M. & Oosterlaan, J. A systematic review and meta-analysis of neuroimaging in Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD) taking Attention-Deficit Hyperactivity Disorder (ADHD) into account. *Neuropsychol. Rev.* **26**, 44–72 (2016).
7. Marsh, A.A., *et al.* Reduced amygdala response to fearful expressions in children and adolescents with callous-unemotional traits and disruptive behavior disorders. *Am. J. Psychiatry* **165**, 712–720 (2008).
8. Decety, J., Michalska, K.J., Akitsuki, Y. & Lahey, B.B. Atypical empathic responses in adolescents with aggressive conduct disorder: a functional MRI investigation. *Biol. Psychol.* **80**, 203–211 (2009).
9. Coccaro, E.F., McCloskey, M.S., Fitzgerald, D.A. & Phan, K.L. Amygdala and orbitofrontal reactivity to social threat in individuals with impulsive aggression. *Biol. Psychiatry* **62**, 168–178 (2007).
10. Aghajani, M., *et al.* Disorganized amygdala networks in conduct-disordered juvenile offenders with callous-unemotional traits. *Biol. Psychiatry* **82**, 283–293 (2017).
11. Ibrahim, K., *et al.* Reduced amygdala-prefrontal functional connectivity in children with autism spectrum disorder and co-occurring disruptive behavior. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **4**, 1031–1041 (2019).
12. Ewbank, M.P., *et al.* Psychopathic traits influence amygdala–anterior cingulate cortex connectivity during facial emotion processing. *Soc. Cogn. Affect. Neurosci.* **13**, 525–534 (2018).
13. Stoddard, J., *et al.* Association of irritability and anxiety with the neural mechanisms of implicit face emotion processing in youths with psychopathology. *JAMA psychiatry* **74**, 95–103 (2017).
14. Kryza-Lacombe, M., Iturri, N., Monk, C.S. & Wiggins, J.L. Face emotion processing in pediatric irritability: neural mechanisms in a sample enriched for irritability with autism spectrum disorder. *J. Am. Acad. Child Adolesc. Psychiatry* (2019).
15. Scheinost, D., *et al.* Ten simple rules for predictive modeling of individual differences in neuroimaging. *Neuroimage* **193**, 35–45 (2019).
16. Yip, S.W., Kiluk, B. & Scheinost, D. Toward addiction prediction: An overview of cross-validated predictive modeling findings and considerations for future neuroimaging research. *BP:CNNI* (2019).
17. Finn, E., Shen, X. & Scheinost, D. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci* **18**, 1664–1671 (2015).
18. Lake, E.M., *et al.* The functional brain organization of an individual allows prediction of measures of social abilities transdiagnostically in autism and attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **86**, 315–326 (2019).

19. Yip, S.W., Scheinost, D., Potenza, M.N. & Carroll, K.M. Connectome-based prediction of cocaine abstinence. *Am. J. Psychiatry* **176**, 156–164 (2019).
20. Shen, X., *et al.* Using connectomebased predictive modeling to predict individual behavior from brain connectivity. *nature protocols. Nat. Protoc.* **12**, 506–518 (2017).
21. Shmueli, G. To Explain or to Predict? *Statistical Science* **25**, 289–310 (2010).
22. Casey, B., *et al.* The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).
23. Hart, H., Radua, J., Nakao, T., Mataix-Cols, D. & Rubia, K. Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. *JAMA psychiatry* **70**, 185–198 (2013).
24. Cohen, A.O., *et al.* When is an adolescent an adult? Assessing cognitive control in emotional and nonemotional contexts. *Psychol. Sci.* **27**, 549–562 (2016).
25. Tottenham, N., *et al.* The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* **168**, 242–249 (2009).
26. Blair, R., Veroude, K. & Buitelaar, J. Neuro-cognitive system dysfunction and symptom sets: a review of fMRI studies in youth with conduct problems. *Neurosci. Biobehav. Rev.* (2016).
27. Raschle, N.M., *et al.* Atypical dorsolateral prefrontal activity in female adolescents with conduct disorder during effortful emotion regulation. *BP:CNNI* **4**, 984–994 (2019).
28. Etkin, A., Egner, T. & Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn. Sci.* **15**, 85–93 (2011).
29. Milad, M.R. & Quirk, G.J. Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* **420**, 70–74 (2002).
30. Silvers, J.A., *et al.* vIPFC–vmPFC–amygdala interactions underlie age-related differences in cognitive regulation of emotion. *Cereb. Cortex* **27**, 3502–3514 (2016).
31. van den Heuvel, M.P. & Sporns, O. A cross-disorder connectome landscape of brain dysconnectivity. *Nature reviews neuroscience* **20**, 435–446 (2019).
32. McRae, K., *et al.* The development of emotion regulation: an fMRI study of cognitive reappraisal in children, adolescents and young adults. *Social cognitive and affective neuroscience* **7**, 11–22 (2012).
33. Buhle, J.T., *et al.* Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. *Cerebral cortex* **24**, 2981–2990 (2014).
34. Ibrahim, K., *et al.* Sex differences in medial prefrontal and parietal cortex structure in children with disruptive behavior. *Dev. Cogn. Neurosci.* **47**, 100884 (2021).
35. Smaragdi, A., *et al.* Sex differences in the relationship between conduct disorder and cortical structure in adolescents. *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 703–712 (2017).
36. Hyatt, C.J., Haney-Caron, E. & Stevens, M.C. Cortical thickness and folding deficits in conduct-disordered adolescents. *Biological Psychiatry* **72**, 207–214 (2012).

37. Jiang, W., *et al.* Reduced cortical thickness and increased surface area in antisocial personality disorder. *Neuroscience* **337**, 143–152 (2016).
38. Buckner, R.L., Andrews-Hanna, J.R. & Schacter, D.L. The brain's default network: anatomy, function, and relevance to disease. (2008).
39. Mars, R.B., *et al.* On the relationship between the “default mode network” and the “social brain”. *Front. Hum. Neurosci.* **6**, 189 (2012).
40. Rosen, M.L., *et al.* Salience network response to changes in emotional expressions of others is heightened during early adolescence: relevance for social functioning. *Developmental Science* **21**, 12571 (2018).
41. Seeley, W.W., *et al.* Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* **27**, 2349–2356 (2007).
42. Chiong, W., *et al.* The salience network causally influences default mode network activity during moral reasoning. *Brain* **136**, 1929–1941 (2013).
43. Aghajani, M., *et al.* Dissociable relations between amygdala subregional networks and psychopathy trait dimensions in conduct-disordered juvenile offenders. *Hum. Brain Mapp.* **37**, 4017–4033 (2016).
44. Werhahn, J.E., *et al.* Aggression subtypes relate to distinct resting state functional connectivity in children and adolescents with disruptive behavior. *Eur. Child Adolesc. Psychiatry*, 1–13 (2020).
45. Broulidakis, M.J., *et al.* Reduced default mode connectivity in adolescents with conduct disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **55**, 800–808 (2016).
46. Cohn, M.D., *et al.* Differential relations between juvenile psychopathic traits and resting state network connectivity. *Hum. Brain Mapp.* **36**, 2396–2405 (2015).
47. Dalwani, M.S., *et al.* Default mode network activity in male adolescents with conduct and substance use disorder. *Drug Alcohol Depend.* **134**, 242–250 (2014).
48. Sebastian, C.L., *et al.* Neural responses to affective and cognitive theory of mind in children with conduct problems and varying levels of callous-unemotional traits. *Arch. Gen. Psychiatry* **69**, 814–822 (2012).
49. Choe, D.E., Shaw, D.S. & Forbes, E.E. Maladaptive social information processing in childhood predicts young men's atypical amygdala reactivity to threat. *J. Child Psychol. Psychiatry* **56**, 549–557 (2015).
50. Capage, L. & Watson, A.C. Individual differences in theory of mind, aggressive behavior, and social skills in young children. *Early Education and Development* **12**, 613–628 (2001).
51. Mandy, W., Skuse, D., Steer, C., St Pourcain, B. & Oliver, B.R. Oppositionality and socioemotional competence: interacting risk factors in the development of childhood conduct disorder symptoms. *J. Am. Acad. Child Adolesc. Psychiatry* **52**, 718–727 (2013).
52. Reeck, C., Ames, D.R. & Ochsner, K.N. The social regulation of emotion: An integrative, cross-disciplinary model. *Trends Cogn. Sci.* **20**, 47–63 (2016).
53. Fairchild, G., Van Goozen, S.H., Calder, A.J., Stollery, S.J. & Goodyer, I.M. Deficits in facial expression recognition in male adolescents with early-onset or adolescence-onset conduct disorder. *J. Child*

- Psychol. Psychiatry* **50**, 627–636 (2009).
54. Lovett, B.J. & Sheffield, R.A. Affective empathy deficits in aggressive children and adolescents: A critical review. *Clin. Psychol. Rev.* **27**, 1–13 (2007).
 55. Chudzik, L. Moral judgment and conduct disorder intensity in adolescents involved in delinquency: Matching controls by school grade. *Psychol. Rep.* **101**, 221–236 (2007).
 56. Finn, E.S., *et al.* Can brain state be manipulated to emphasize individual differences in functional connectivity? *Neuroimage* **160**, 140–151 (2017).
 57. Yarkoni, T. & Westfall, J. Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspect. Psychol. Sci.* **12**, 1100–1122 (2017).
 58. Marek, S., *et al.* Towards reproducible brain-wide association studies. *bioRxiv* (2020).
 59. Sha, Z., Wager, T.D., Mechelli, A. & He, Y. Common dysfunction of large-scale neurocognitive networks across psychiatric disorders. *Biol. Psychiatry* **85**, 379–388 (2019).
 60. Cremers, H.R., Wager, T.D. & Yarkoni, T. The relation between statistical power and inference in fMRI. *PLoS One* **12**, 0184923 (2017).
 61. Sukhodolsky, D.G., *et al.* Neural mechanisms of cognitive-behavioral therapy for aggression in children and adolescents: design of a randomized controlled trial within the National Institute for Mental Health Research Domain Criteria construct of frustrative non-reward. *J. Child Adolesc. Psychopharmacol.* **26**, 38–48 (2016).
 62. Achenbach, T.M. & Rescorla, L.A. *Manual for the ASEBA School-Age Forms & Profiles* (University of Vermont, Research Center for Children, Youth, and Families, Burlington, VT, 2001).
 63. Kaufman, J., Birmaher, B., Axelson, D., Perepletchikova, F., Brent, D., & Ryan, N. Schedule for Affective Disorders and Schizophrenia for School Aged Children: Present and Lifetime Version for DSM-5 (K-SADS-PL). Retrieved from <https://www.pediatricbipolar.pitt.edu/resources/instruments>. (2016).
 64. Le Couteur, A., Lord, C. & Rutter, M. *The Autism Diagnostic Interview-Revised (ADI-R)*. (Western Psychological Services, CA, Los Angeles, 2003).
 65. Lord, C., Rutter, M., DiLavore, P.C. & Risi, S. *Autism Diagnostic Observation Schedule—Second Edition (ADOS-2)* (Western Psychological Services, Los Angeles, CA, 2012).
 66. Raine, A., *et al.* The reactive–proactive aggression questionnaire: differential correlates of reactive and proactive aggression in adolescent boys. *Aggress. Behav.* **32**, 159–171 (2006).
 67. Bushman, B.J. & Anderson, C.A. Is it time to pull the plug on hostile versus instrumental aggression dichotomy? *Psychol. Rev.* **108**, 273 (2001).
 68. Smeets, K.C., *et al.* Are proactive and reactive aggression meaningful distinctions in adolescents? A variable-and person-based approach. *J. Abnorm. Child Psychol.* **45**, 1–14 (2017).
 69. Ibrahim, K., *et al.* Anger rumination is associated with restricted and repetitive behaviors in children with autism spectrum disorder. *J. Autism Dev. Disord.* **49**, 3656–3668 (2019).
 70. Wechsler, D. *WAIS-III Administration and Scoring Manual* (The Psychological Corporation, San Antonio, TX, 1997).

71. Constantino, J.N. *Social Responsiveness Scale (SRS)* (Western Psychological Services, Los Angeles, 2005).
72. Cholemkey, H., Kitzerow, J., Rohrmann, S. & Freitag, C.M. Validity of the social responsiveness scale to differentiate between autism spectrum disorders and disruptive behaviour disorders. *Eur. Child Adolesc. Psychiatry* **23**, 81–93 (2014).
73. Crick, N.R. & Dodge, K.A. A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychol. Bull.* **115**, 74 (1994).
74. Dodge, K.A. & Coie, J.D. Social-information-processing factors in reactive and proactive aggression in children's peer groups. *J. Pers. Soc. Psychol.* **53**, 1146 (1987).
75. Helseth, S.A., Waschbusch, D.A., King, S. & Willoughby, M.T. Aggression in children with conduct problems and callous-unemotional traits: Social information processing and response to peer provocation. *J. Abnorm. Child Psychol.* **43**, 1503–1514 (2015).
76. Smith, S.M., *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23**, S208-S219 (2004).
77. Woolrich, M.W., *et al.* Bayesian analysis of neuroimaging data in FSL. *Neuroimage* **45**, S173-S186 (2009).
78. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825–841 (2002).
79. Greene, A.S., Gao, S., Scheinost, D. & Constable, R.T. Task-induced brain state manipulation improves prediction of individual traits. *Nature communications* **9**, 1–13 (2018).
80. Noble, S., *et al.* Influences on the test–retest reliability of functional connectivity MRI and its relationship with behavioral utility. *Cerebral Cortex* **27**, 5415–5429 (2017).
81. Rapuano, K.M., *et al.* Behavioral and brain signatures of substance use vulnerability in childhood. *Dev. Cogn. Neurosci.* **46**, 100878 (2020).
82. Whelan, R., *et al.* Adolescent impulsivity phenotypes characterized by distinct brain networks. *Nat. Neurosci.* **15**, 920 (2012).

Table

Table 1. Participant Demographics and Clinical Characteristics

Variable	Total Sample N=129	Aggressive Behavior n = 100	Healthy Controls n=29	p value ^b
Age, years (SD)	11.9 (2.2)	11.7 (2.3)	12.8 (1.8)	
Male, %	69	73	55.2	0.02 ^c
Mean IQ ^a (SD)	107.8 (13.8)	106.4 (14.1)	112.5 (12.3)	0.07
Race, %				0.04 ^c
White	76	78	69	0.45
Black	13.2	12	17.2	
American Indian/ Alaska native	1.6	2	0	
Asian/Pacific Islander	1.6	2	0	
Other/More than one race	7.8	6	13.8	
Ethnicity				1
Hispanic	15.5	16	13.8	
Non-Hispanic	84.5	84	86.2	
Mean CBCL aggressive behavior T score (SD)	69.8 (12.2)	75.3 (7.6)	51 (2.7)	<0.001 ^d
Mean CBCL internalizing behavior T score (SD)	58.6 (13.1)	63.3 (10.3)	42.4 (7.3)	<0.001 ^d
Mean SRS-2 SCI total T score (SD)	60.7 (13.8)	65.5 (11.3)	44.1 (7.8)	<0.001 ^d
RPQ Aggression Total (SD)	15.7 (9.3)	19.4 (6.9)	2.8 (3.1)	
DSM-5 diagnoses ^e , %				
Oppositional defiant disorder		76		
Conduct disorder		4		
DBD-NOS		3		
DMDD		18		
ASD		18		
ADHD		78		
Anxiety disorder		26		
Depression disorder		4		
Type of medication, %				
Stimulants		31		
Antidepressant		13		
Neuroleptics		13		
Non-stimulants		20		
Mood stabilizers		4		
Benzodiazepines		2		

Note: Diagnoses of disruptive behavior disorders and comorbid disorders were made using the *Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version* (K-SADS-PL). Following DSM-5, oppositional defiant disorder diagnosis was not assigned to children who met criteria for DMDD.

Abbreviations: ASD, autism spectrum disorder; ADHD, Attention-deficit/hyperactivity disorder; CBCL, Child Behavior Checklist; DMDD, Disruptive mood dysregulation disorder; HC, healthy controls; RPQ, Reactive-Proactive Anger Questionnaire; SRS-2 SCI, Social Responsiveness Scale-Second Edition Social Communication and Interaction.

^aFull-scale IQ measured by the Wechsler Abbreviated Scale of Intelligence⁷⁰.

^bSignificant group differences at $p < 0.05$ using Chi-square test for categorical variables or independent samples T-test.

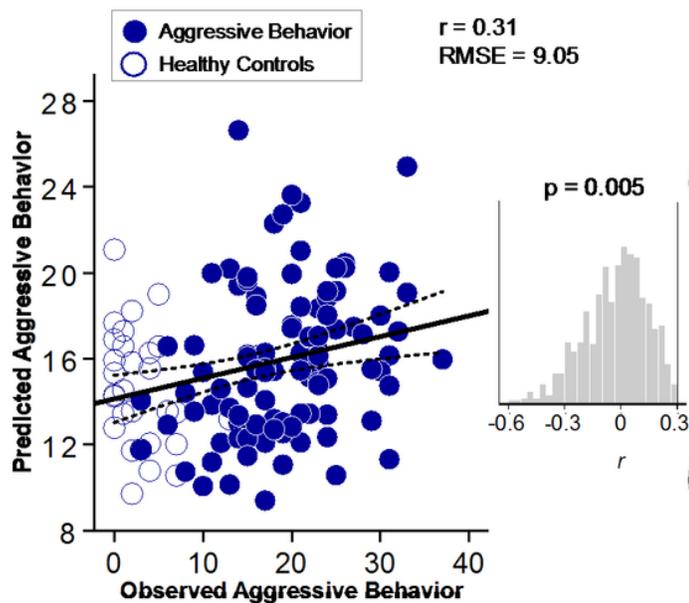
^cHC > Aggressive Behavior group

^dAggressive Behavior group > HC

^eFollowing DSM-5, oppositional defiant disorder diagnosis was not assigned to children who met criteria for DMDD.

Figures

a Model Fit: Predicted vs. Observed



b Positive and Negative Networks Predicting Aggression

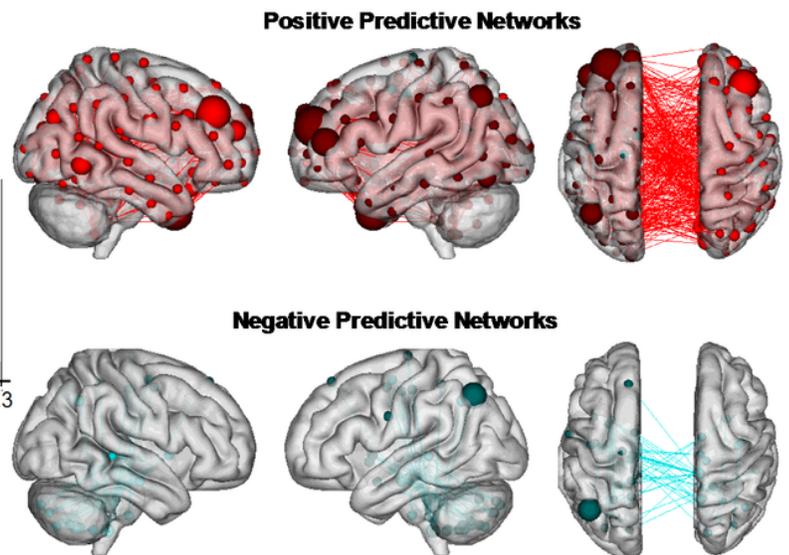
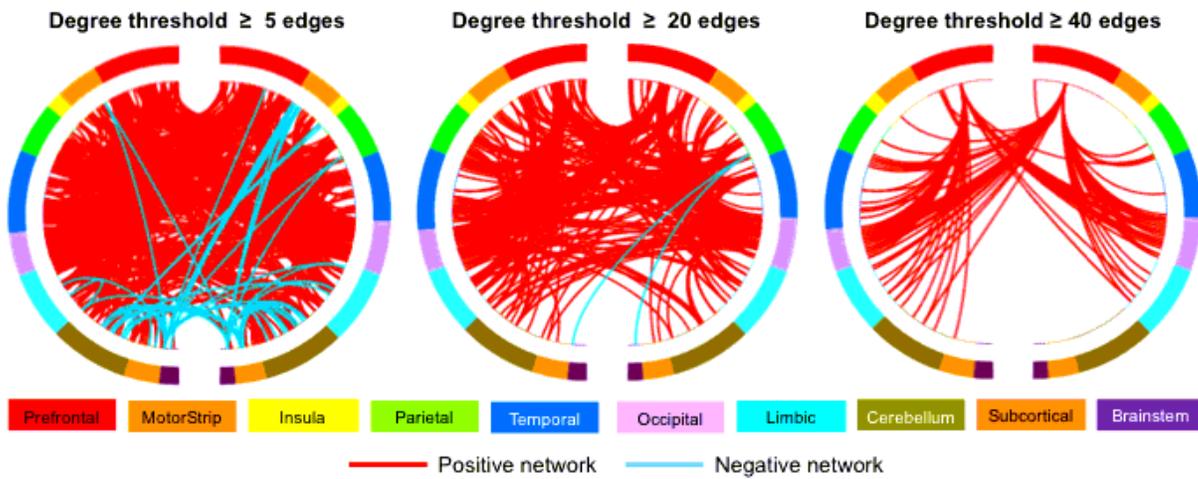


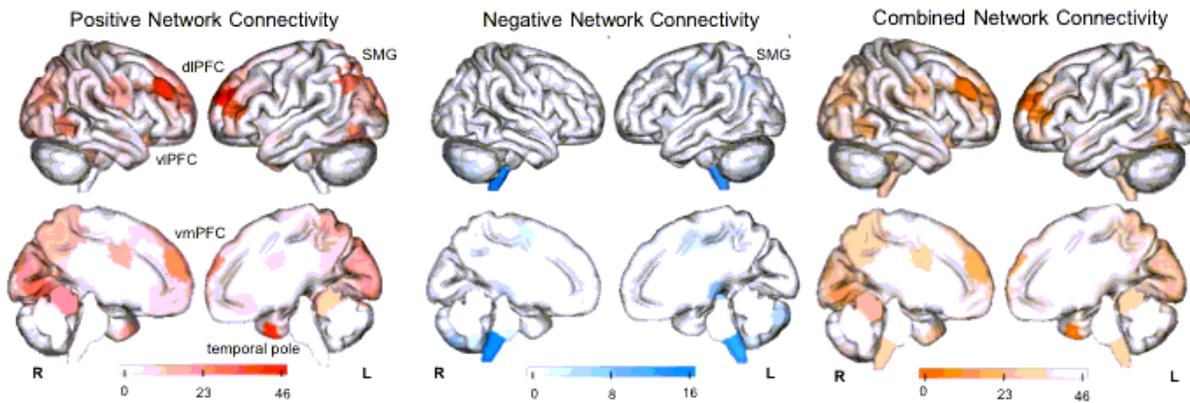
Figure 1

Brain-wide functional connectivity predicts severity of aggressive behavior. (a) Panel A illustrates the correspondence between observed (x-axis) and predicted (y-axis) severity of aggressive behavior generated using CPM. The Reactive-Proactive Aggression Questionnaire (RPQ) was used as a continuous measure of severity of aggressive behavior. Despite the clinical complexity of the population, CPM successfully predicted aggression ($p = 0.005$, permutation testing). Predictions remained significant in follow-up analyses controlling for covariates including age, IQ, sex, medication use as well as co-occurring ADHD and internalizing symptoms. Histogram shows distribution of Pearson correlation (r) values from 1000 iterations of randomly-shuffled ratings of aggression severity used to nonparametrically determine P values. (b) Panel B shows positive (red) and negative (blue) networks predicting aggression. For positive networks, increased edge weights (i.e., increased functional connectivity) predict greater severity of aggression. For negative networks, decreased edge weights (i.e., decreased functional connectivity) predict greater severity of aggression. Larger spheres indicate nodes with more edges, and smaller spheres indicate fewer edges. Note: RMSE = Root Mean Square Error.

a Edges that contributed to the CPM model



b Connectivity nodes



c Within and between network functional connectivity

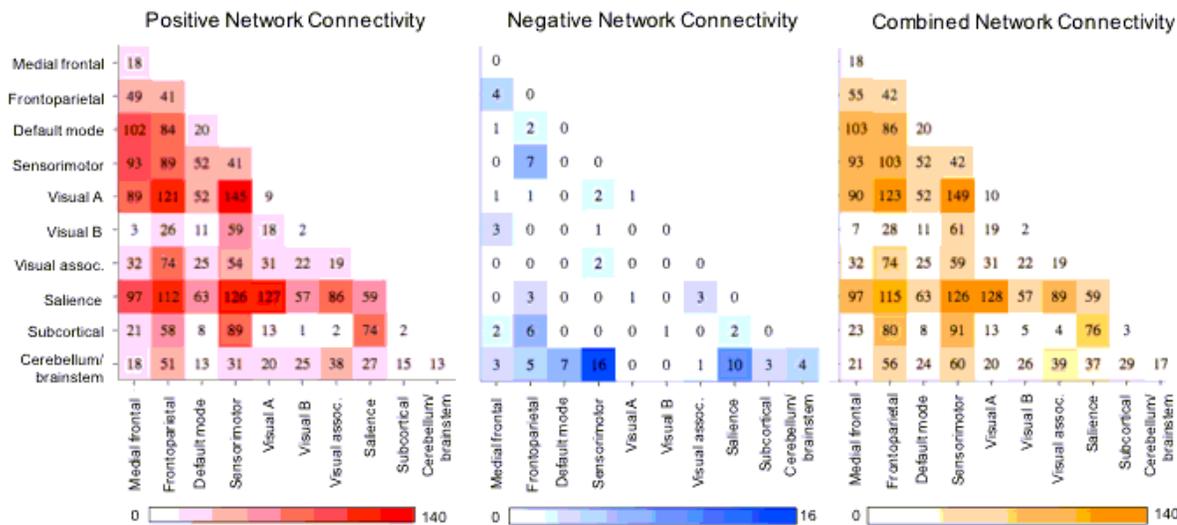


Figure 2

Networks predicting aggression summarized by connectivity between macroscale brain regions and networks. (a) From the top, brain regions are presented in approximate anatomical order whereby longer-range connections are represented by longer lines. To visualize these complex networks and clarify edges that contributed the most to the CPM model, edges belonging to nodes with five or more edges (degree ≥ 5 ; left), 20 or more edges (degree ≥ 20 ; middle), and 40 or more edges (degree ≥ 40 ; right) are shown. (b)

Visualization of node degree (i.e., the sum of predictive edges for a node) for the positive (left), negative networks (middle), and all networks (right). Darker colors indicate higher degree nodes (i.e., with more edges) contributing to CPM model. Several lateral and ventral prefrontal nodes emerged as high degree positive predictive nodes including bilateral dorsolateral prefrontal cortex (dlPFC), ventrolateral and ventromedial PFC (vlPFC and vmPFC, respectively). A negative predictive node emerged in the supramarginal gyrus (SMG) as well as sensorimotor regions. (c) Within- and between-network connectivity for the positive network (left), negative network (middle), and combined networks (right) are summarized using canonical networks. Cells represent the total number of edges connecting nodes within and between each network, with darker colors indicating a greater number of edges. By definition, positive and negative networks do not contain overlapping edges.

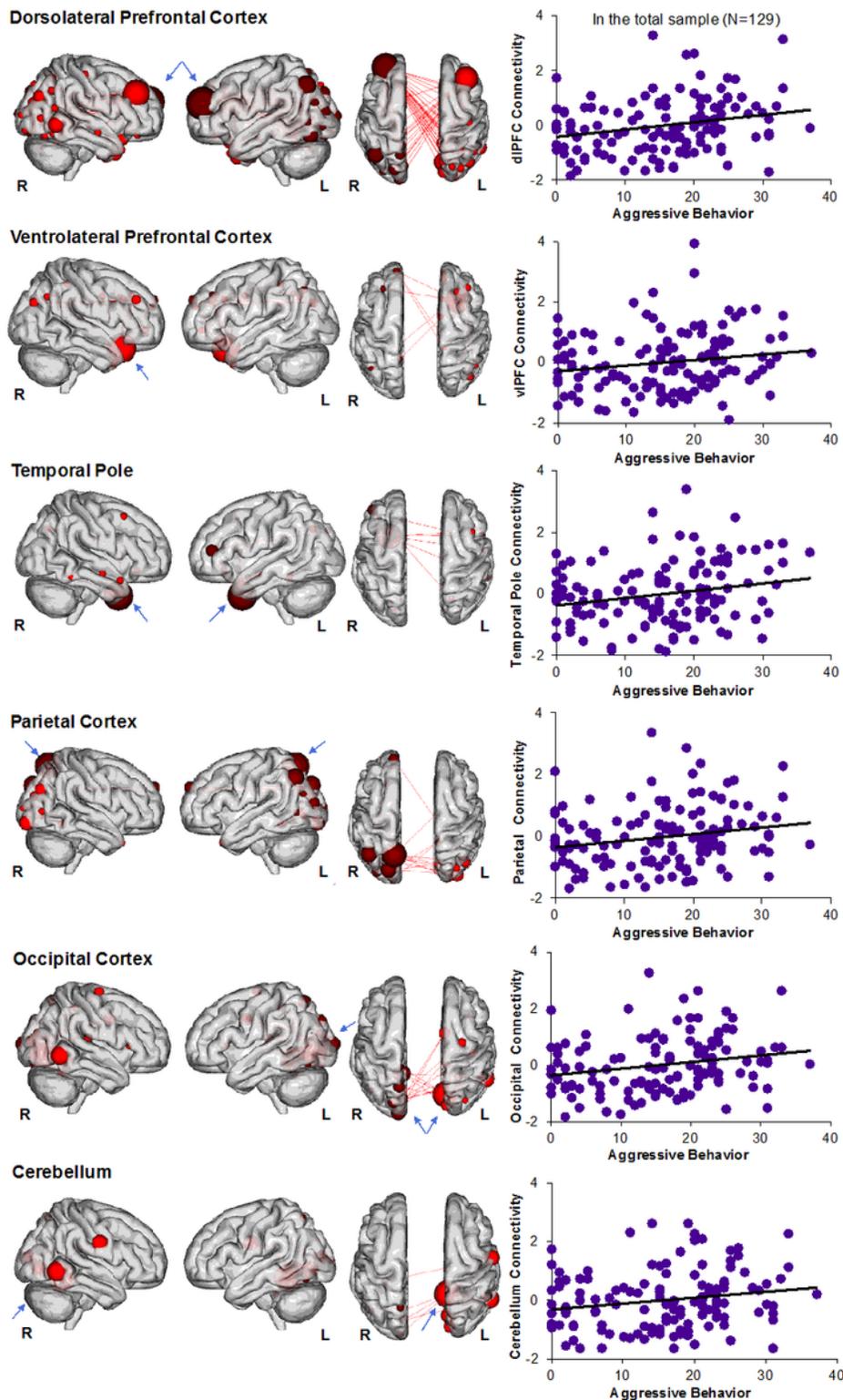


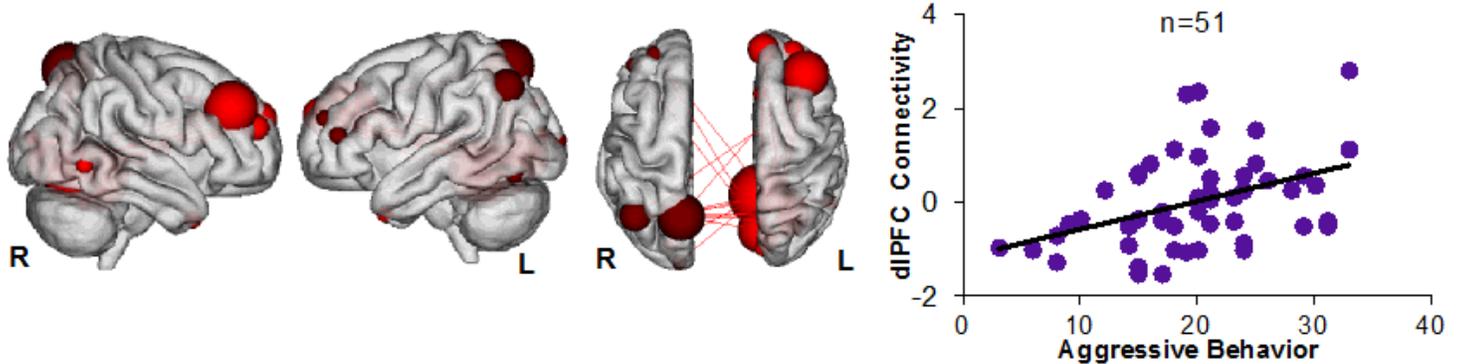
Figure 3

Follow up analyses for high-degree nodes contributing to the connectome model in the discovery sample of 129 children. Follow up analyses tested the sensitivity of high-degree nodes (i.e., most predictive features) in predicting aggressive behavior. Regions emerging as highly predictive of aggression in the connectome model include the bilateral dorsolateral prefrontal cortex (dIPFC), right ventrolateral PFC (vIPFC), bilateral temporal pole, bilateral parietal cortex, bilateral occipital cortex, and right cerebellum.

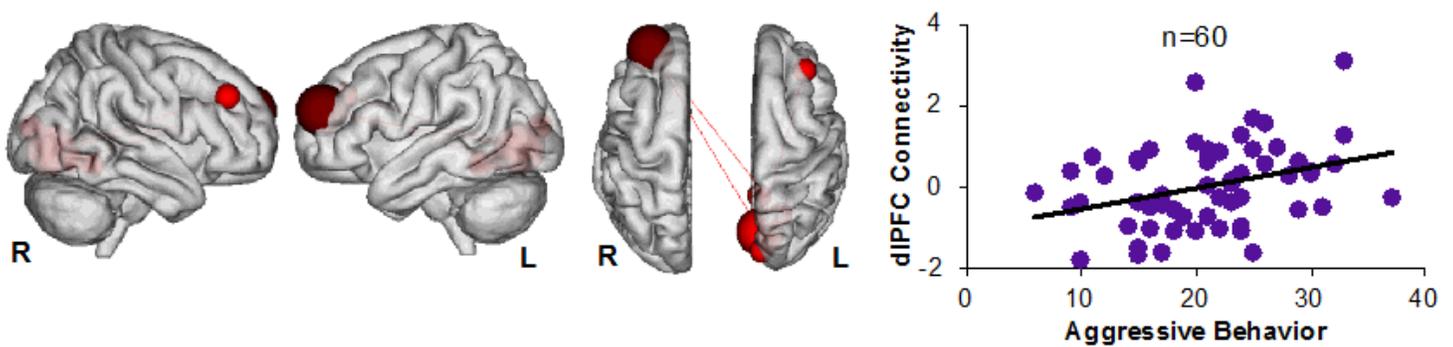
See Table S3 for MNI coordinates of high-degree nodes. Post-hoc tests retained the high-degree nodes and all edges connected to these nodes, while removing all other edges. For illustrative purposes, scatter plots depict the strength of node connectivity plotted on the y-axis and severity of aggressive behavior (Reactive-Proactive Aggression Questionnaire total) on the x-axis.

Dorsolateral Prefrontal Cortex Connectivity

a Subgroup with high severity of co-occurring internalizing symptoms



b Subgroup with high severity of co-occurring ADHD symptoms



c Subgroup with high severity of co-occurring social behavior impairments

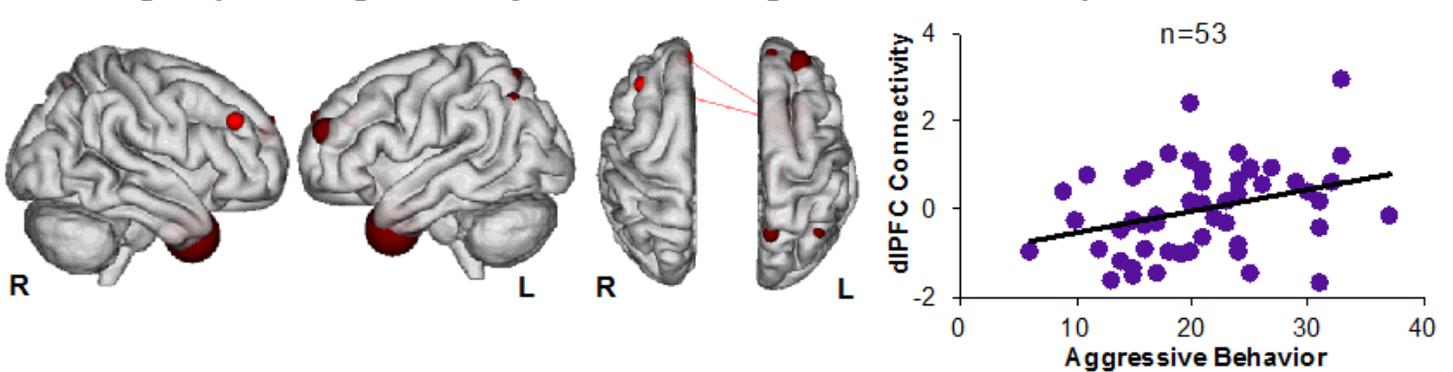


Figure 4

Dorsolateral prefrontal cortex connectivity predicts aggressive behavior. Two nodes in the bilateral dorsolateral prefrontal cortex (dIPFC) emerged as the highest degree nodes (i.e., most predictive) in the CPM model predicting aggression (MNI coordinates: left dIPFC, $x=-28, y=50, z=21$; right dIPFC, $x=37, y=35, z=31$). Larger spheres indicate nodes with more edges, and smaller spheres indicate fewer edges. a-c Follow up analyses tested the sensitivity of these dIPFC networks in predicting aggressive behavior

despite commonly co-occurring behaviors with aggression. From the total sample, subgroups were formed to test the robustness of the dlPFC networks in predicting aggression based on cutpoints using a T score >65, which represents a clinical threshold on standardized measures (CBCL, SRS-2 SCI): (a) high severity of internalizing symptoms (n=51); (b) high severity of ADHD symptoms (n=60); and (c) high severity of social behavior impairments (n=53). Post-hoc tests retained the bilateral dlPFC nodes and all edges connected to these nodes, while removing all other edges. For illustrative purposes, scatter plots depict the strength of dlPFC connectivity plotted on the y-axis and severity of aggressive behavior (Reactive-Proactive Aggression Questionnaire total) on the x-axis.

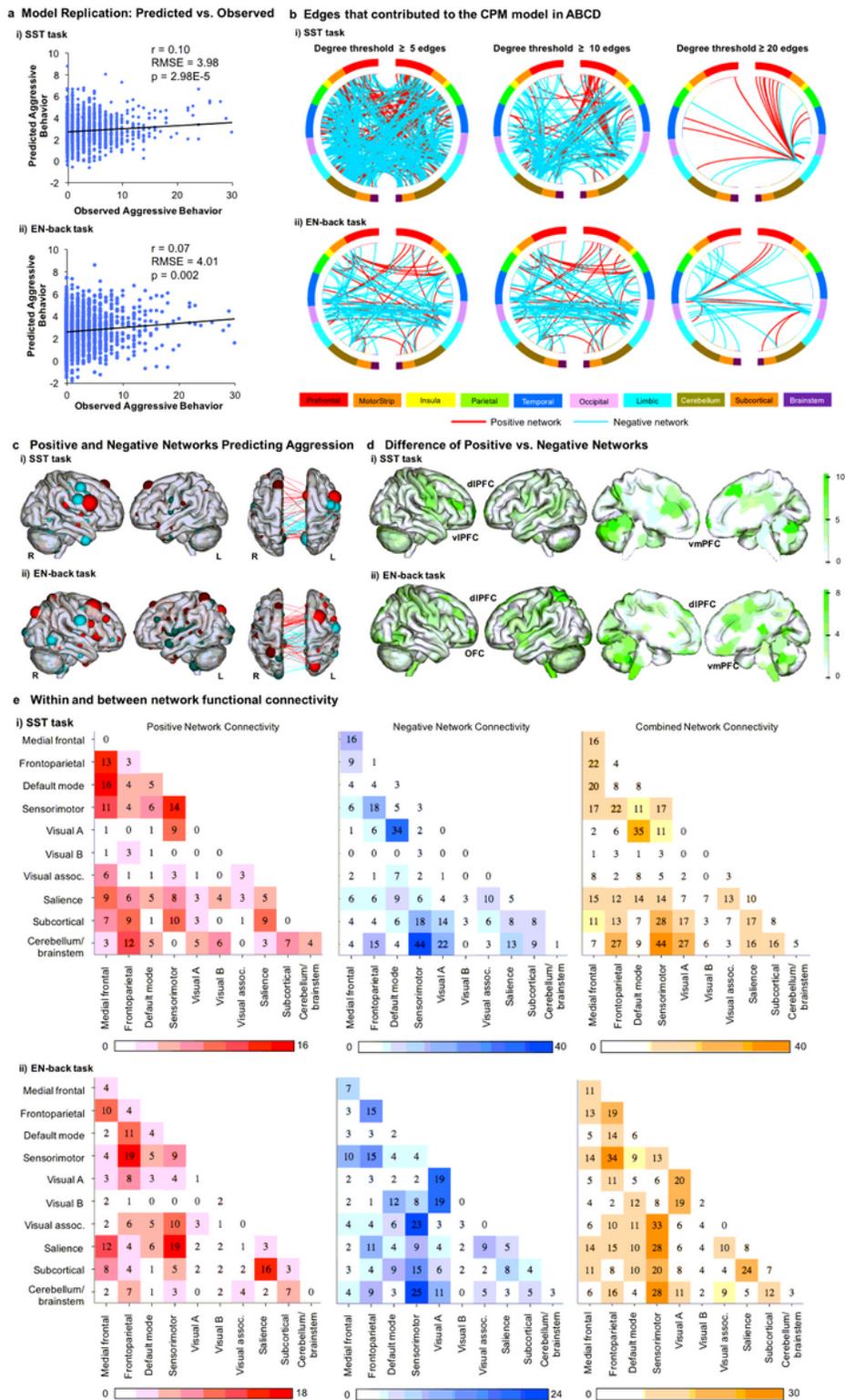


Figure 5

Replication of findings using CPM prediction of aggression in an out-of-sample dataset. a-e Using the Adolescent Brain Cognitive Development (ABCD) study dataset, we trained and tested a separate CPM model predicting aggressive behavior from 1701 children (920 females) using the stop signal task (SST) and from 1791 children (958 females) using the emotional n-back task (EN-back) (age range 9-10 years). We selected these tasks for replication and external validation because the SST and EN-back tasks tap

frontoparietal and fronto-amygdala circuitry relevant to aggressive behavior^{22-24, 82} and related constructs of emotion processing and response inhibition that are implicated in disruptive behavior disorders in children. ABCD data was processed in a similar manner to the main dataset⁸¹. (a) Panel A illustrates the correspondence between observed (x-axis) and predicted (y-axis) severity of aggressive behavior generated using CPM for the SST and EN-back tasks. The CBCL aggressive behavior total score was used as a continuous measure of aggression. (b) Brain regions are presented in approximate anatomical order whereby longer-range connections are represented by longer lines. To visualize these complex networks and clarify edges that contributed the most to the CPM model, edges belonging to nodes with five or more edges (degree ≥ 5 ; left), 10 or more edges (degree ≥ 10 ; middle), and 20 or more edges (degree ≥ 20 ; right) are shown. Warmer colors indicate positive networks and cooler colors indicate negative networks. (c) Several prefrontal regions emerged as high degree predictive nodes including dorsolateral prefrontal cortex (dlPFC), ventrolateral and ventromedial PFC (vlPFC and vmPFC, respectively), and lateral orbitofrontal PFC (OFC). Larger spheres indicate nodes with more edges, and smaller spheres indicate fewer edges. (d) Visualization of node degree (i.e., the sum of predictive edges for a node) for the positive versus negative networks. Darker colors indicate higher positive degree nodes (i.e., with more edges) contributing to CPM model. (e) Within- and between-network connectivity for the positive network (left), negative network (middle), and combined networks (right) are summarized using canonical networks for the SST and EN-back tasks. Cells represent the total number of edges connecting nodes within and between each network. Here, darker colors indicate a greater number of edges. Positive and negative networks do not contain overlapping edges. Note: RMSE = Root Mean Square Error.

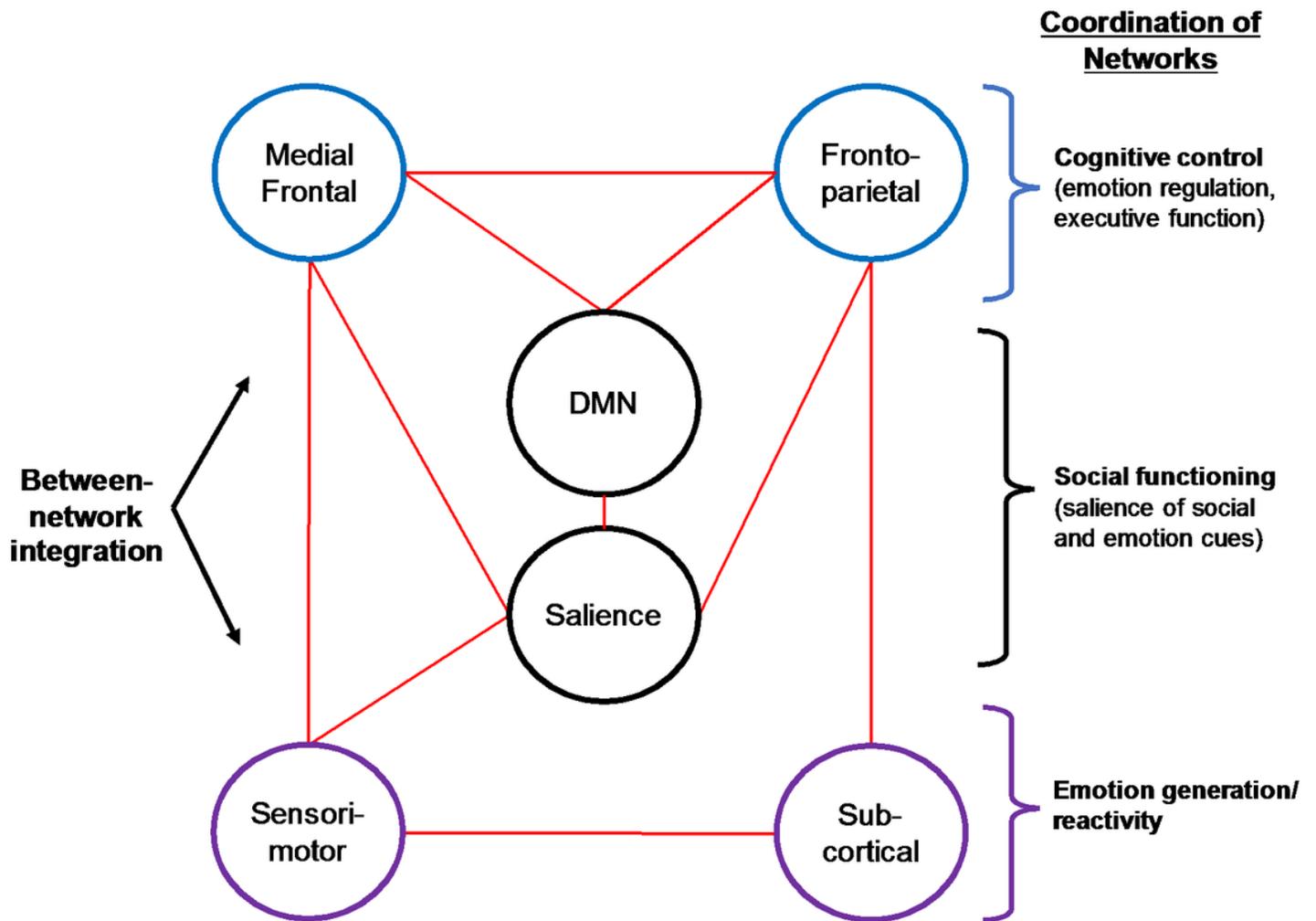


Figure 6

Network model of aggression. Large-scale patterns of between-network connectivity in aggression networks identified using connectome-based predictive modeling are summarized based on relative number of connections within networks. Disrupted connectivity (i.e., network coordination) between frontoparietal and medial-frontal networks (top), between salience and default mode (DMN) networks (middle), and between sensorimotor and subcortical networks (bottom) positively predicted severity of aggressive behavior in children. Disrupted long-range connectivity (i.e., network integration) between networks are indicated in red. To facilitate interpretation, predictive networks are described according to domains of impairments implicated in aggression: cognitive control, social behavior, and emotion generation/reactivity.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AggressionconnectomeNNSupplement20210322NEW8881.docx](#)