Structural brain measures among children with and without ADHD in the Adolescent Brain and Cognitive Development Study cohort: a cross-sectional US population-based study

Joel Bernanke, Alex Luna, Le Chang, Elizabeth Bruno, Jordan Dworkin, Jonathan Posner

Lancet Psychiatry 2022; 9: 222–31

Published Online February 7, 2022 https://doi.org/10.1016/S2215-0366(21)00505-8

See Comment page 187

New York State Psychiatric Institute, New York, NY, USA (J Bernanke MD, A Luna MD, L Chang MS, E Bruno BA, J Dworkin PhD); **Department of Psychiatry, Vagelos College of Physicians and Surgeons** (J Bernanke, A Luna, L Chang, E Bruno, J Dworkin) **and Department of Biostatistics, Mailman School of Public Health** (J Dworkin), **Columbia University, New York, NY, USA; Department of Psychiatry and Behavioral Health, Duke University, Durham, NC, USA** (J Posner MD)

Correspondence to: Dr Jonathan Posner, Duke University, North Pavilion Building, 2400 Pratt Street, Durham, NC 27705, USA **Jonathan.posner@duke.edu**

Summary

Background Structural neuroimaging research has identified a variety of abnormalities in cortical and subcortical structures in children with ADHD. However, studies to date have not employed large, non-referred samples, complete with data on potential confounding variables. Here, we tested for differences in structural MRI measures among children with and without ADHD using data from the Adolescent Brain and Cognitive Development (ABCD) Study, the largest paediatric brain imaging study in the USA.

Methods In this cross-sectional study, we used baseline demographic, clinical, and neuroimaging data from the ABCD Study, which recruited children aged 9–10 years between Sept 1, 2016, and Aug 31, 2018, representative of the sociodemographic features of the US population. ADHD was diagnosed by parent report of symptoms. Neuroimaging data underwent centralised quality control and processing by the ABCD team. Linear mixed effects models were used to estimate Cohen's d values associated with ADHD for 79 brain measures of cortical thickness, cortical area, and subcortical volume. We used a novel simulation strategy to assess the ability to detect significant effects despite potential diagnostic misclassification.

Findings Our sample included 10 736 participants (5592 boys, 5139 girls; 5692 White, 2165 Hispanic, 1543 Black, 221 Asian, and 1100 of other race or ethnicity), of whom, 949 met the criteria for ADHD and 9787 did not. In the full model, which included potential confounding variables selected a priori, we found only 11 significant differences across the 79 brain measures after false discovery rate correction, all indicating reductions in brain measures among participants with ADHD. Cohen's d values were small, ranging from –0·11 to –0·06, and were not meaningfully changed by using a more restrictive comparison group or alternative diagnostic methods. Simulations indicated adequate statistical power to detect differences even if there was substantial diagnostic misclassification.

Interpretation In a sample representative of the general population, children aged 9–10 years with ADHD differed only modestly on structural brain measures from their unaffected peers. Future studies might need to incorporate other MRI modalities, novel statistical approaches, or alternative diagnostic classifications, particularly for research aimed at developing ADHD diagnostic biomarkers.

Funding Edwin S Webster Foundation and Duke University, NC, USA.

Introduction

ADHD is a debilitating neurodevelopmental condition marked by difficulties in one or both of the following domains: sustained concentration or hyperactivity and impulsivity. In epidemiological studies in the USA, ADHD affects about 9% of children;[1,2] worldwide estimates are slightly lower.[3,4] ADHD is associated with increased risk for disruptive behaviour disorders, mood or anxiety disorders, and substance use disorders.[5–8]

Structural neuroimaging studies have attempted to identify differences in brain measures among children with ADHD compared to controls. Small studies have shown moderate effect sizes in frontal and subcortical structures[9,10] but underpowered studies risk producing false-positive results.[11] Larger studies have generally reported smaller effect sizes, though they have still found a variety of structural differences.[12–14] These differences include cortical thinning in the frontal, parietal, and occipital lobes, but especially in prefrontal regions; reduced subcortical volumes, particularly in the basal ganglia; and reduced cerebellum and total brain volume.[15]

Notable in this literature are two mega-analyses by the ENIGMA ADHD Working Group. The first focused on subcortical structures and reported reduced volumes in the accumbens, amygdala, caudate, hippocampus, and putamen, as well as reduced estimated total intracranial volume (ETIV) among children with ADHD.[16] Statistically significant Cohen's d values were between –0·19 and –0·13, with the number of cases and controls ranging from 604 to 854. The ENIGMA group's second mega-analysis of cortical measures showed differences in surface area across 24 regions and in cortical thickness across four regions among children.[17] In this study, significant Cohen's d values ranged from –0·21 to –0·10,

Research in context

Evidence before this study

Many structural neuroimaging studies have attempted to identify differences in brain measures among children with ADHD compared with their peers. We searched PubMed using the terms (Attention Deficit Hyperactivity Disorder [Title] OR ADHD [Title]) AND (meta-analysis [Title/Abstract] OR mega-analysis [Title/Abstract]) AND (structural [Title/Abstract] OR MRI [Title/Abstract]) AND English [Language] and found nine meta-analyses and two mega-analyses published before Aug 1, 2021. Meta-analyses have confirmed a variety of differences, including, but not limited to, cortical thinning in the frontal, parietal, and occipital lobes, especially in prefrontal regions; reduced subcortical volumes, particularly in the basal ganglia; and reduced cerebellum and total brain volume. Two mega-analyses by the ENIGMA ADHD Working Group (Hoogman et al, 2017; Hoogman et al, 2019) showed a large number of statistically significant differences in brain measures, with effect sizes ranging from –0·19 to –0·13 in the study of subcortical volumes, and –0·21 to –0·10 in the study of cortical structures. Limitations of existing ADHD neuroimaging studies include the use of convenience, clinical, or referred samples that do not represent the general population, as well as the limited availability of data on potential confounding variables.

Added value of this study

To our knowledge, this is both the largest structural neuroimaging study of ADHD to date and the first to use a non-referred, national sample. We were able to control for several potential confounding variables, including participants' age, sex, socioeconomic status, race or ethnicity, and comorbidities, such as anxiety, obsessive-compulsive disorder, and depressive disorders. Our results suggest that in the general population, children aged 9–10 years with ADHD differ only modestly from their unaffected peers on structural MRI measures.

Implications of all the available evidence

We observed fewer differences between children with ADHD and their unaffected peers than previous research, which might be explained by the cohort in our study being older, having fewer severe cases of ADHD, or being more heterogeneous than the samples used in other studies. Future studies might need to incorporate other MRI modalities, novel statistical approaches, or alternative diagnostic classifications, particularly for research aimed at developing ADHD diagnostic biomarkers.

with the numbers of cases and controls ranging from 974 to 1081. In both studies, younger children had generally larger and more extensive differences than older children and adolescents. All ENIGMA analyses were adjusted for age, sex, site, and, when indicated, ETIV. The authors found little to no influence of intelligence, comorbidity, medication use, or factors related to image acquisition and processing.

One limitation of existing ADHD neuroimaging studies is their use of convenience, clinical, or referred samples that do not represent the general population. For example, in the ENIGMA study of cortical measures, only 16% of children had diagnoses of oppositional defiant disorder (ODD) and 8·6% had anxiety disorders, which is lower than the rates of comorbidity in clinical and epidemiological samples.[6,8] Another limitation has been incomplete data on potential confounding variables. In the ENIGMA studies, only a subset of participants had any confounder data, and socioeconomic and pubertal measures were absent.

The development of large datasets, such as the Adolescent Brain and Cognitive Development (ABCD) Study, offers new opportunities to explore the structural correlates of ADHD. The ABCD Study is the largest longitudinal study of brain development in the USA. The recruitment strategy of this study, including the locations of the 22 participating sites, was designed to obtain a locally random cohort of children aged 9–10 years, representative of the sociodemographic features of the population in the USA.[18,19] In our study, we compare structural brain measures among children with and without ADHD in the ABCD sample.

A common criticism of large, non-clinical studies, such as ABCD, is that their greater sample sizes can come at the expense of careful clinical phenotyping, which might lead to increased rates of diagnostic misclassification compared with smaller, disorder-specific studies, resulting in falsely diminished effect sizes. To address this, we used a novel simulation strategy to assess our power to detect significant effects despite misclassification.

Methods

Participants

Participants in the ABCD Study were recruited through the school system, with school selection based on demographic factors. Participants were included if they were in the required age range (aged 9–10 years), and were able to provide informed consent (parents) and assent (child). Participants were excluded if they were not proficient in English, or if parents were not fluent in English or Spanish; had a severe sensory, medical, or neurological condition that would make their data invalid or limit their ability to follow the study protocols; or had contraindications to undergoing MRI.[20] The number of school systems are not described in the ABCD dataset. The accumulating sample was monitored, and recruitment adjusted, to reach pre-selected targets.[19] Recruitment occurred between Sept 1, 2016, and Aug 31, 2018, although the study is ongoing. Recruitment protocols have been published previously.[19]

We conducted analyses on the ABCD Study 2.0.1 data release, which contained baseline assessments and imaging data for all 11875 participants, including 2100 twins, and 30 triplets. Institutional review boards at each site approved the study procedures. Written consent was obtained from all parents and verbal assent was given by all children.

Structural neuroimaging

Participants underwent structural imaging on 3T MRI platforms to obtain high resolution T1-weighted images (1 mm isotropic). Acquisition parameters are described elsewhere.[18] Structural data were processed by the ABCD team with FreeSurfer (version 5.3.0) per standardised processing pipelines, including skull stripping, white matter segmentation, correction of topological defects, surface optimisation, and non-linear registration to a spherical surface-based atlas. In our analyses, we used the 35 regions labelled with the Desikan-Killiany atlas for cortical thickness and area measures; the eight subcortical volumes identified by the FreeSurfer pipeline; and ETIV, for a total of 79 brain measures per participant. All these regions were the same regions used in the ENIGMA studies.

See Online for appendix

The ABCD team performed quality control, manually reviewing T1-weighted images and FreeSurfer cortical surface reconstructions. Software was used to estimate and score head motion (0=absent, 1=mild, 2=moderate, and 3=severe). We excluded any imaging results the ABCD team marked as unacceptable (eg, gross imaging artifacts).[21]

Demographic, neurocognitive, and clinical measures

All ABCD Study participants underwent a battery of psychometric and neuroimaging assessments.[22] Socio-demographic data were collected from caregivers. Participants from the same family were identified accordingly. Children and caregivers completed the pubertal development scale, but due to greater missingness in the children's reports, we used the caregiver's score on this measure. Children underwent neurocognitive testing using the National Institutes of Health (NIH) Toolbox.[23] We used the age-corrected composite score, a standardised, normed estimate of intelligence comparable to other intelligence quotient (IQ) measures.

Outcomes

The paper-and-pencil Kiddie-Schedule for Affective Disorders and Schizophrenia (KSADS) is often considered the gold-standard in psychiatric diagnosis among youth. The KSADS has been used as the diagnostic instrument in numerous NIH and industry-sponsored trials and has been translated into over 30 languages. The computerised-KSADS (KSADS-COMP) retains the core elements of, and shows agreement with, the clinician-administered paper-and-pencil version. It is available in English, Spanish, and Danish. In a validation study, most parents and children completed the self-administered KSADS-COMP in 90 min or less.

For more on **KSADS-COMP** see https://www.kennedykrieger.org/ksads-comp

We chose to use the KSADS-COMP categorical diagnosis of ADHD as the primary outcome because it is available in the ABCD dataset; is consistent with the DSM-5 criteria; is based on a well-studied and validated tool, both in research and clinic settings, including epidemiological studies; and has favourable validity data. We did not distinguish between ADHD presentations.

For secondary outcomes, we used the Child Behavior Checklist (CBCL)-parent report, and the Brief Problem Monitor (BPM)-teacher report. We defined a combined anxiety and depression variable from the KSADS-COMP, which was positive if criteria were met for one or more anxiety (including obsessive compulsive) or depressive disorders (appendix p 2). All parent report measures were completed by a single caregiver; the dataset does not indicate if the caregiver was the mother, father, or some other adult.

Data cleaning, imputation, and linear mixed effects modelling

Data cleaning, imputation, and analyses were performed in R (version 4.0.2). The entire ABCD 2.0.1 release was loaded using the pre-made R-data file available from the ABCD team. The dataset was then limited to baseline observations. Participants missing KSADS-COMP ADHD diagnoses were removed. We compared the characteristics of participants with ADHD (sex ratio and prevalence of various comorbidities) to estimates from previous epidemiological studies in the USA. Missing values for all covariates except for ADHD, study site, scanner, and family were imputed and summary statistics for the original and imputed datasets were compared. To be consistent with the simulation study (described later), a single merged, imputed dataset was used for primary analyses. Parameter estimates and p values obtained using Rubin's rules for combining results from imputed datasets were essentially identical to those obtained using this single, merged, imputed dataset (appendix p 2).

FreeSurfer outputs were averaged across the right and left hemispheres. We used linear mixed effects models to estimate the association between ADHD and brain measures using the lme4 package (version 1.1.2.3). The modelling output for two sets of fixed effect terms was compared. In the full model set, we included covariates selected a priori that would probably confound the relationship between ADHD and brain measures, including age (months), sex assigned at birth (male, female), race or ethnicity (White, Black, Hispanic, Asian, other), pubertal development (1–4), estimated IQ, comorbidity (anxiety or depression), and motion during scanning (0–3); and three indicators of socioeconomic status, parental education (did not complete high school, high school diploma or General Educational Development

test, some college, undergraduate degree, graduate degree), parental income (<US$50 000, $50 000–$100 000, ≥$100 000), and parental marital status (yes or no). The ABCD Study collected a variety of potential socioeconomic indicators. We focused on these three as they are commonly used to operationalise socioeconomic status, are more objective than some indicators, and are largely present in the ABCD dataset. In the reduced model set, we included only age and sex (the covariates used in the ENIGMA analyses). In both sets, we included ETIV as a fixed effects term for cortical area and subcortical volume measures, but not for cortical thickness measures. In neither set did we include ODD as a covariate, because it was strongly associated with ADHD.

In all models, we included family nested in scanner as an intercept-varying random effects term to account for differences across scanners (some sites had more than one) and similarities within families. Modelling did not distinguish between twins, triplets, and other siblings. All variables were treated as unordered factors, except age, NIH Toolbox score, and ETIV, which were treated as numeric. Cohen's d values were calculated manually from modelling output (appendix pp 2–3).[24] To account for multiple comparisons, false discovery rate (FDR) correction was performed separately for the 35 cortical area measures, 35 cortical thickness measures, and eight subcortical volume measures plus ETIV.[25] Corrected p values <0·05 were considered significant.

We tested the possibility that the associations between ADHD and brain measures were different among boys and girls, and among participants with and without comorbidities, by adding ADHD-by-sex and ADHD-by-comorbidity interaction terms to the full model. For ADHD and comorbidity, we used the combined comorbidity variable described earlier. FDR correction was performed on the p values for the interaction terms. We also tested whether associations were different across racial and ethnic groups by adding an ADHD-by-race or ethnicity interaction term to the full model. Because there are more than two racial or ethnic groups, we used a likelihood ratio test to determine if adding this interaction term improved the model, and again used FDR correction to adjust the p values. All FDR corrections were performed separately for cortical area, cortical thickness, and subcortical volume measures. We calculated sex-specific, comorbidity-specific, and race or ethnicity-specific effect sizes for brain measures with interaction term or likelihood ratio test p values <0·05.

Alternative diagnostic classifications

To test whether our results were affected by the choice of comparison group or diagnostic instrument, we performed a series of alternative analyses. First, we excluded participants with major psychiatric conditions from the comparison group to form a healthy comparison group. Excluded conditions were major depression, bipolar disorder, generalised anxiety disorder, obsessive-compulsive disorder, anorexia, bulimia, and post-traumatic stress disorder, as determined by the KSADS-COMP parent report. Second, we used the CBCL to diagnose ADHD, marking children with T-scores of 65 or more on the DSM-5 ADHD subscale as having the disorder. This subscale is more accurate at diagnosing ADHD than the Attention Problems subscale.[26] Although T-scores of 65–69 are considered borderline, a cutoff of 65 or more is more sensitive and almost as specific among children in community settings.[27]

Third, we used two restrictive, multi-informant approaches to diagnose ADHD: (1) having the KSADS-COMP parent report diagnosis plus a T-score of 65 or more on the BPM teacher report Attention Problems subscale, and (2) having T-scores of 65 or more on both the relevant CBCL (parent) and BPM (teacher) subscales. To account for differences in sample size and composition between the subset of participants with BPM data and the full dataset, we also calculated effect sizes using only the KSADS-COMP and CBCL diagnoses for just the participants with BPM data. We then compared summary measures for the effect sizes (mean and SD) generated using the KSADS-COMP, CBCL, KSADS-COMP plus BPM, and CBCL plus BPM diagnostic criteria for the subset of participants with BPM data. Finally, we calculated the phi correlation coefficient and intraclass correlation coefficient (ICC) for all four methods of diagnosing ADHD.

Simulations

To assess the potential effect of ADHD misclassification, we simulated new brain measurements for each participant in the ABCD dataset that reflected the effect sizes reported in the ENIGMA studies. We then tested whether various amounts of misclassification could obscure ENIGMA-like effect sizes in the ABCD sample. Specifically, we used three misclassification regimes: (1) 25% of ADHD diagnoses were false positives; (2) 10% of no ADHD diagnoses were false negatives; and (3) a combination of both. We ran the simulation 1000 times for each regime. With each run, we randomly selected participants to reclassify, repeated our mixed effects modelling using the ENIGMA covariates, and counted the number of statistically significant brain differences (appendix p 3). We report the distribution of significant results for the three misclassification regimes, and a no misclassification regime, and compare them to the number of significant effects observed in the ABCD sample.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

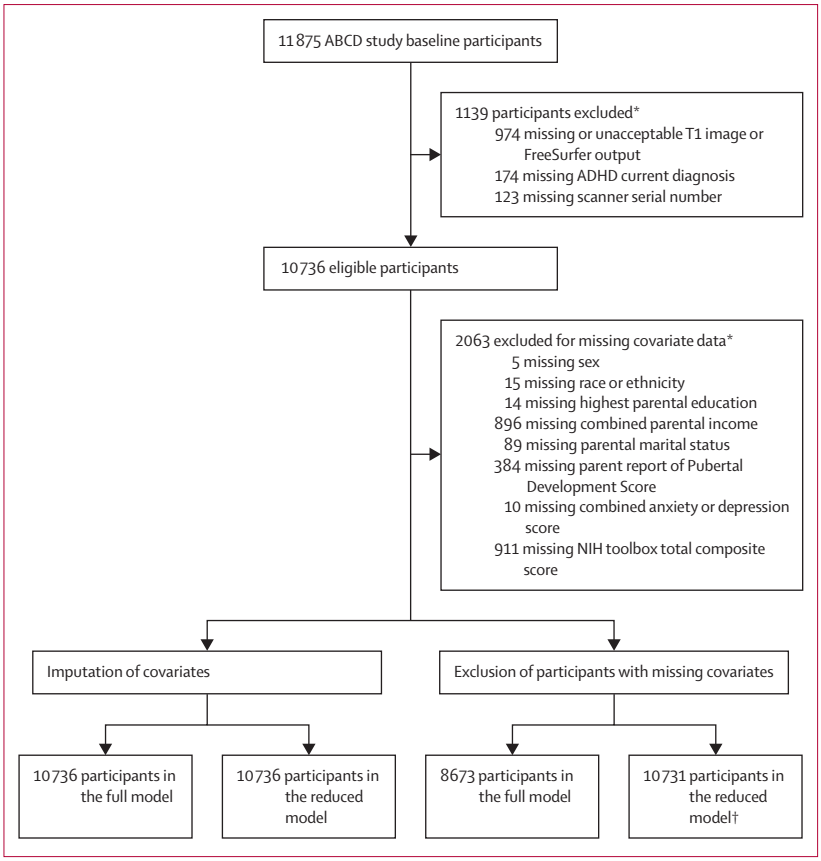

Figure 1: **Flow diagram of study participants**
ABCD=Adolescent Brain and Cognitive Development. NIH=National Institutes of Health. *Some participants were missing more than one variable. †This group is all eligible participants minus those with missing sex data.

Results

The original ABCD dataset included 11 875 participants. After exclusion of individuals missing ADHD diagnoses or with unacceptable imaging data, 10 736 remained. Exclusion of participants who were missing sex data reduced the sample to 10 731, and exclusion of participants missing any covariate data reduced the sample to 8673. Imputing covariates made all 10 736 eligible participants available for analysis (figure 1). Imputation did not alter the distributions of the original variables.

Descriptive statistics for the 10 736 eligible participants are given in table 1. Of these, 949 (8·9%) participants met the criteria for ADHD on the KSADS-COMP parent report. A greater percentage of participants with ADHD were boys than girls (2·21:1, 68·8% *vs* 50·5%, p<0·0001). Participants with ADHD had higher prevalence of comorbid ODD (29·8% *vs* 3·2%, p<0·0001) and any anxiety or depressive disorder than participants without ADHD (40·7% *vs* 13·1%, p<0·0001) consistent with patterns of comorbidity reported in community and clinic samples.[5–8] The ADHD group had a slightly lower mean NIH Toolbox composite score (44·7 *vs* 48·2, p<0·0001) and pubertal development score (1·6 *vs* 1·8), and slightly higher levels of head motion during MRI scanning (mean 0·7 *vs* 0·6) than participants without ADHD. There were small but statistically significant differences in demographic and socioeconomic indicators, including the distribution of racial or ethnic groups, parental education, and parental marital status.

Using the merged, imputed dataset, and the full mixed effects model, we found 11 significant differences in brain measures (among the 79 we tested) between children with and without ADHD. All the effects indicated reductions in measures in children with ADHD relative to the comparison group. Ten out of the 11 differences were cortical area measures; the 11th was ETIV. No significant differences were seen in cortical thickness measures or in subcortical volumes. Significant effect sizes ranged from –0·11 to –0·06. The largest differences were seen in the caudal anterior cingulate (d=–0·11, FDR p=0·013) and pericalcarine cortex (d=–0·11, FDR p=0·018; appendix pp 5–7). Limiting modelling to the 8673 participants (772 with ADHD, 7901 without ADHD) without any missing covariate data produced comparable effect sizes, but only ETIV remained significant. None of the 11 measures showed significant ADHD-by-sex interactions, although two out of the full set of 79 brain measures did (appendix p 8). No brain measures had significant ADHD-by-comorbidity interactions, and none of the likelihood ratio tests indicated improved model fit with an ADHD-by-race or ethnicity model term.

The reduced mixed effects model, with only age and sex as covariates, yielded a similar result to the full model, identifying nine significant differences, all also identified by the full model; only the lingual gyrus (d=–0·07, FDR p=0·084) and precentral gyrus (d=–0·06, FDR p=0·066) were no longer significant. Significant effect sizes in the reduced model ranged from –0·12 to –0·06. The caudal anterior cingulate (d=–0·11, FDR p=0·013) and pericalcarine cortex (d=–0·11, FDR p=0·013) again showed the largest differences, along with ETIV (d=–0·12, FDR p=0·0003; appendix pp 10–12). The omission of the five participants without sex data did not alter these results. The results for the full and reduced models, along with the results from the ENIGMA studies, are shown in figure 2.

Of the 9787 participants without ADHD in the main analysis, 1004 had a major psychiatric condition and were excluded, leaving 8783 to form a healthy comparison group. Effect sizes estimated using the full set of covariates were essentially unchanged, although only eight remained significant (appendix pp 14–16). Seven of the 10 736 participants in the main analysis were missing CBCL subscale scores. Using the ADHD subscale T-score cutoff of 65 or more, 660 participants met the criteria for ADHD and 10 069 did not, with some participants moving from the ADHD to the comparison group, and vice versa. Phi correlation coefficient and ICC between the KSADS-COMP and CBCL diagnoses indicated a strong positive relationship and moderate

reliability, respectively (appendix p 21). Effect sizes, again estimated using the full set of covariates, were similar to the original results, but the largest and only significant effect was for ETIV (d=–0·10, FDR p=0·027; appendix pp 18–20).

Only 3178 participants (29·6%) had BPM teacher report data. Of those, 95 (3·0%) had both a KSADS-COMP diagnosis of ADHD and a BPM T-score of 65 or more, compared with 283 (8·9%) who had a KSADS-COMP diagnosis of ADHD (regardless of BPM T-score). 72 (2·3%) participants had both a CBCL and BPM T-score of 65 or more, compared with 187 who had a CBCL T-score of 65 or more (regardless of BPM T-score). We used the full model to estimate the effect sizes associated with these four different methods of diagnosing ADHD. None of the resulting differences were statistically significant, though effect sizes were larger when using the most restrictive diagnostic method (table 2). Phi correlation coefficient and ICC among all four diagnostic methods ranged from moderate to strong or good agreement (appendix p 21).

In the no misclassification simulation regime, the minimum number of significant results, assuming the ENIGMA effect sizes were accurate, was 41 (mean 51·8). In the misclassification regimes, a false-positive rate of 25% resulted in 237 (25%, rounding to the nearest person) of 949 participants diagnosed with ADHD being reclassified as not having the condition. A false-negative rate of 10% meant 979 (10%, rounding to the nearest person) of 9787 participants without ADHD were reclassified as having it, more than were diagnosed with ADHD initially. These reclassifications produced a prevalence of 16%, sensitivity of 42%, and specificity of 97% in the combined regime. The minimum number of significant results under the false negative only regime was 36 (mean 48·1), false positive only regime was 28 (mean 40·7), and combined regime was 22 (mean 34·2). Distributions for all 1000 simulations for each regime are shown in figure 3.

Discussion

Our study is the largest structural neuroimaging study of ADHD to date. We compared cortical surface area, cortical thickness, and subcortical volume measures among children aged 9–10 years with and without ADHD. By contrast with previous meta-analyses and mega-analyses, which have shown effect sizes in the small to moderate range across numerous brain regions, our results showed both diminished effect sizes and a reduced number of statistically significant differences, in accordance with an overall trend towards smaller effect sizes as study samples become larger.

In line with most structural studies of ADHD, our statistically significant results all indicated reductions in brain measures among children with ADHD. Similar to the ENGIMA study of cortical regions, the differences we detected were concentrated in surface area measures, including total surface area.[17] Consistent with prior meta-analyses, mega-analyses, or reviews, we detected differences in the anterior and posterior cingulate cortex,

	ADHD (N=949)	No ADHD (N=9787)	p value
Age, months	118·6 (7·4)	119·1 (7·5)	0·075
Sex	..	..	<0·0001
Assigned female at birth	296 (31·2%)	4843 (49·5%)	..
Assigned male at birth	653 (68·8%)	4939 (50·5%)	..
Missing	0	5 (<1%)	..
Race or ethnicity	..	..	<0·0001
White	527 (55·5%)	5165 (52·8%)	..
Black	151 (15·9%)	1392 (14·2%)	..
Hispanic	148 (15·6%)	2017 (20·6%)	..
Asian	5 (<1%)	216 (2·2%)	..
Other	117 (12·3%)	983 (10·0%)	..
Missing	1 (<1%)	14 (<1%)	..
Highest parental education	..	..	<0·0001
No high school diploma	25 (2·6%)	476 (4·9%)	..
High school diploma or general educational diploma	66 (7·0%)	920 (9·4%)	..
Some college	284 (29·9%)	2489 (25·4%)	..
Bachelor's degree	265 (27·9%)	2479 (25·3%)	..
Graduate degree	308 (32·5%)	3410 (34·8%)	..
Missing	1 (<1%)	13 (<1%)	..
Income	..	..	0·21
<US$50 000 per year	273 (31·3%)	2573 (26·3%)	..
$50 000–$100 000 per year	250 (26·3%)	2562 (26·2%)	..
>$100 000 per year	350 (36·9%)	3832 (39·2%)	..
Missing	76 (8·0%)	820 (8·4%)	..
Parents married	614 (64·7%)	6696 (68·4%)	0·020
Missing	8 (<1%)	81 (<1%)	..
Oppositional defiant disorder	283 (29·8%)	309 (3·2%)	<0·0001
Missing	0	0	..
Anxiety or depression	386 (40·7%)	1285 (13·1%)	<0·0001
Missing	0	10 (<1%)	..
CBCL ADHD DSM-5 subscale	63·5 (7·7)	52·1 (4·1)	<0·0001
Missing	0	7 (<1%)	..
BPM-T attention subscale	61·2 (7·9)	54·4 (6·2)	<0·0001
Missing	666 (70·2%)	6892 (70·4%)	..
PDS-P	1·6 (0·8)	1·8 (0·9)	0·0001
Missing	32 (3·4%)	352 (3·6%)	..
NIH Toolbox Score	44·7 (11·3)	48·2 (11·1)	<0·0001
Missing	79 (8·3%)	832 (8·5%)	..
Motion score	..	..	0·0012
0, absent	384 (40·5%)	4299 (43·9%)	..
1, mild	469 (49·4%)	4822 (49·3%)	..
2, moderate	95 (10·0%)	662 (6·8%)	..
3, severe	1 (0·1%)	4 (0·0%)	..
Missing	0	0	..

Data are mean (SD) or n (%). Means were compared using two sample t-tests. Proportions were compared using χ^2 tests or Fisher's exact test. BPM-T=Brief problem monitor, teacher report. CBCL=Child behavior checklist. PDS-P=pubertal development score, parent report. NIH=National Institutes of Health.

Table 1: **Characteristics of study participants.**

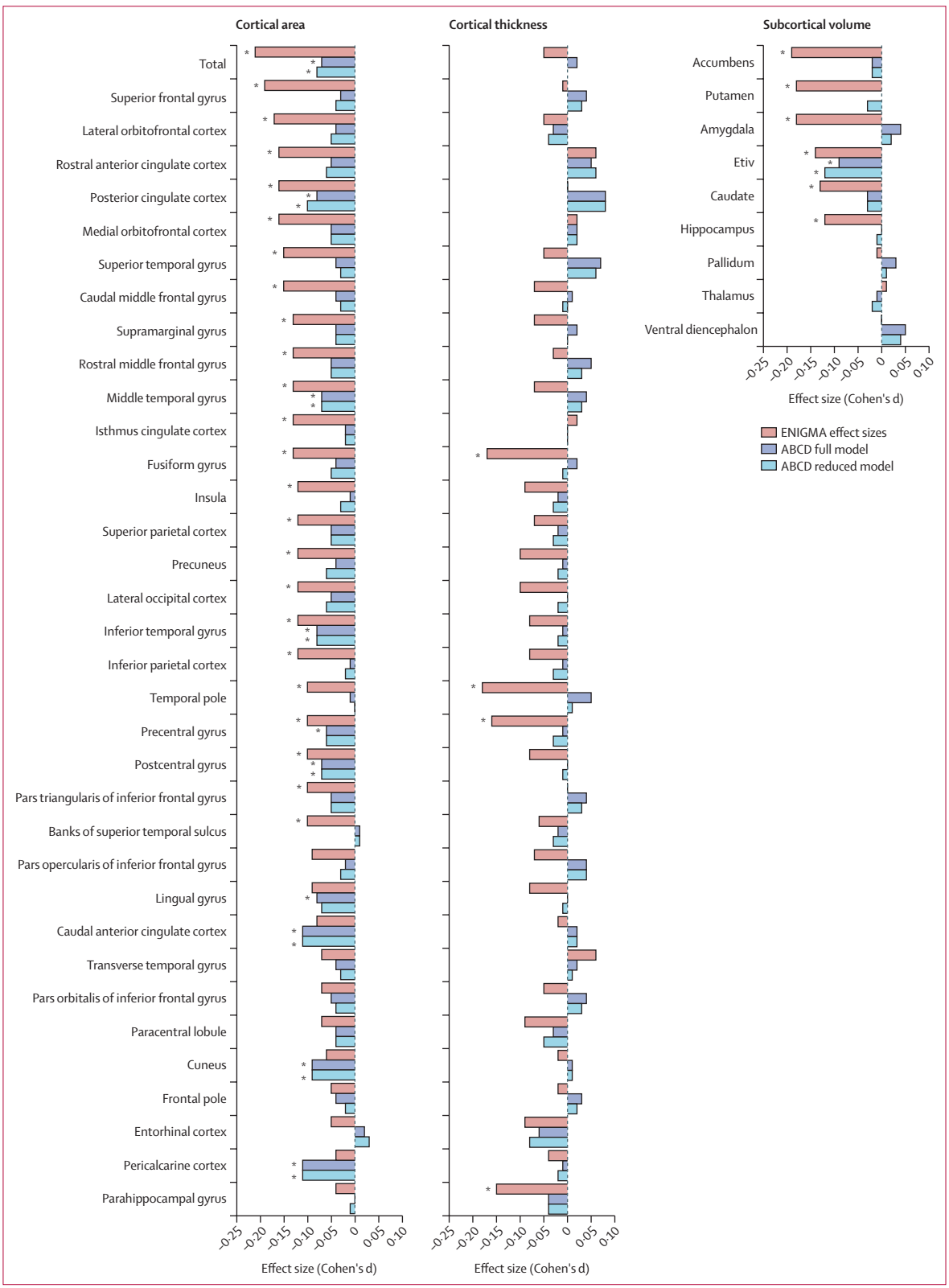

Figure 2: **Effect sizes among children in the ENIGMA study and estimated effect sizes using the full and reduced models in the ABCD cohort**
ABCD=Adolescent Brain and Cognitive Development. ETIV=estimated total intracranial volume. *Significant effect size (false discovery rate p<0·05).

precentral and postcentral gyri, and middle temporal gyrus.[13,17] These regions have been implicated in salience detection, behavioural inhibition, and sensory integration (domains important to the ADHD phenotype).[28] Other differences we found, such as those in the pericalcarine cortex and cuneus, have been seen in individual structural studies, or implicated in functional MRI studies, related to impulsivity and ADHD.[13,29]

By contrast with the ENIGMA and other ADHD studies of subcortical structures, we did not identify differences in subcortical volumes.[16,30] Although there is substantial overlap in our findings and brain areas implicated by previous research, most prior work showed volumetric, rather than surface area, differences. Some of our findings, such as reduced total surface area, reduced ETIV, and changes in the inferior temporal gyrus and the cuneus, have also been seen in one or more other psychiatric conditions, including psychosis, bipolar disorder, and depression.[31–33] Finally, one brain area identified in our study, the lingual gyrus, has primarily been implicated in anxiety and depression, not ADHD.[34]

Possible explanations for the reduced magnitude and number of differences in our study versus the existing literature include methodological choices and misdiagnosis. Several factors, however, make these explanations unlikely. First, we used the ABCD Study's ready-made dataset, which employed a unified scanning protocol, centralised data processing, and standardised quality control. Additionally, our results do not appear to depend on our statistical procedures. Using imputation and additional covariates increased the number of significant differences compared with using the original data or just the ENIGMA covariates, with no meaningful change in effect sizes in either case. Second, the diagnosis of ADHD using the KDADS-COMP produced a case population in the ABCD cohort whose prevalence, sex ratio, and comorbidity profile were comparable with epidemiological studies using other diagnostic tools.[6] Third, excluding participants with major psychiatric conditions from the comparison group, and using the CBCL to redefine the case and comparison groups, did not increase the number of significant findings or the magnitude of the effect sizes.

Lastly, our simulations suggest that diagnostic misclassification was unlikely to obscure ENIGMA-like effect sizes in the ABCD sample, with the minimum number significant findings totalling 22 in the combined regime, twice as many as we observed. While we do not know the true extent of misclassification, our combined regime looks robust. The 25% false-positive rate and 10% false-negative rate of the combined regime resulted in an ADHD prevalence of 16%, sensitivity of 42%, and specificity of 97%. Although 97% specificity might seem high, this figure is an artifact of the ratio of controls to cases. To further reduce specificity, we would have to assume an even higher prevalence of ADHD or an even lower sensitivity. The true prevalence being greater than 16%, or the sensitivity of the KSADS-COMP being below 42%, seems improbable.

More probable explanations for the divergence from previous results include our use of a large sample, the severity of ADHD symptoms in our sample, the neurobiological heterogeneity of ADHD, and the ages of the ABCD Study participants.[35] The large sample size in our study reduces the risk of false-positive findings that might arise from small, non-random sampling from a heterogeneous population of cases and controls, raising the possibility that our results more accurately capture differences (or the absence thereof) in brain measures linked with ADHD in the general population. However, our use of an epidemiological sample also probably produced a case population with less severe symptoms than those seen in the clinical or referred samples

	ADHD	No ADHD	Mean (SD) of effect sizes	Number of brain regions with statistically significant effect sizes
KSADS-COMP (all participants)	949	9787	–0·02 (0·04)	11/79
KSADS-COMP*	283	2895	–0·04 (0·06)	0/79
KSADS-COMP and BPM	95	3083	–0·06 (0·08)	0/79
CBCL*	187	3991	–0·04 (0·06)	0/79
CBCL and BPM	72	3106	–0·11 (0·09)	0/79

Data are n or n/N unless otherwise specified. False discovery rate p values were non-significant in all cases (p>0·05). BPM=Brief Problem Monitor, teacher report. CBCL=Child Behavior Checklist. KSADS-COMP=computerised Kiddie Schedule for Affective Disorders and Schizophrenia for School-aged Children, parent report. *Participants with BPM data.

Table 2: **Full model summary measures for different ADHD diagnostic methods**

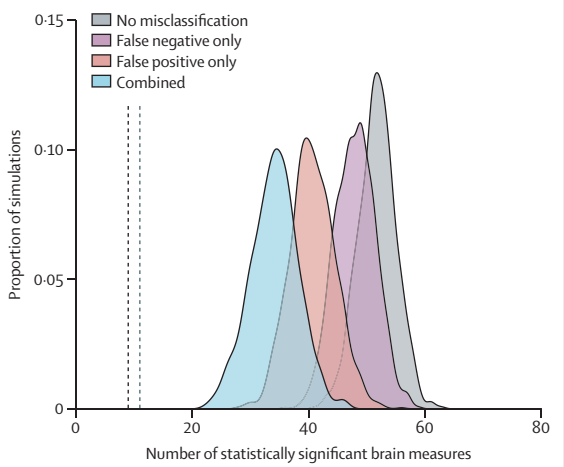

Figure 3: **Density plots showing simulated brain measurements for participants in the ABCD dataset using three misclassification regimes**
These plots show the distribution of significant findings (false discovery rate p<0·05) under the three misclassification regimes (25% of ADHD diagnoses were false positives [red], 10% of no ADHD diagnoses were false negatives [purple], and a combination of both [blue]) and the no misclassification regime (green), for all 1000 runs of each simulation. The vertical dashed line indicates the number of significant findings using the reduced model (nine) and the vertical dotted line indicates the number using the full model (11).

featured in prior research, and some studies have found a relationship between symptom severity and the magnitude of brain differences.[36] Consistent with this interpretation, our analysis using more restrictive and multi-informant diagnostic criteria, potentially identifying more symptomatic cases, yielded effect sizes generally larger than those seen using the KSADS-COMP-parent report criteria.

Another possible explanation for our results is that previous studies, with more restrictive inclusion and exclusion criteria, could have selected more homogeneous subgroups of children with ADHD who have larger structural differences compared with their non-affected peers. Some studies have found associations between biological factors such as age, sex, and long-term medication use, and the degree or presence of particular structural differences.[10,37] Finally, longitudinal studies and cross-sectional studies across a range of ages, including the ENIGMA studies, have found that structural brain differences diminish during childhood, making our results potentially consistent with studies that showed more extensive structural differences before age nine years.[38]

The present study has several limitations. We focused exclusively on structural measures; our results have no bearing on functional or white matter differences. Our study relied on linear modelling; more complex relationships between ADHD and differences in brain measures might require more nuanced statistical approaches. There is no diagnostic validation study nested with the ABCD Study to provide a clear estimate of the accuracy of the KSADS-COMP. Limited information on ADHD-specific treatments prevents us from investigating their possible effects on brain measures. The high prevalence of ODD among participants with ADHD precludes identifying ADHD effects independent of ODD. Additionally, subdividing cases by potential modifying variables runs the same risks associated with smaller samples, limiting our ability to identify brain differences between subgroups.

In view of our findings, future studies should explore alternative ways of defining ADHD to produce more homogeneous case populations, and non-linear approaches. As an alternative to a categorical definition, researchers should consider ADHD-associated traits, such as reward valuation and response inhibition.[39] Alternatively, researchers might consider novel diagnostic approaches based on new, data-driven nosologies.[40] All these approaches probably require better powered and more diverse samples (eg, race or ethnicity and age).

Our results suggest that structural brain differences in a general sample of youth in the USA are smaller in scope and magnitude than previously reported in clinical populations. Future studies might need to incorporate other MRI modalities, novel statistical approaches, or alternative diagnostic classifications, particularly for research aimed at developing ADHD diagnostic biomarkers.

Contributors

All authors had full access to the data in the study. JB, AL, LC, JD, and JP verified the data. JB, AL, LC, JD, and JP conceived and designed the study. JB, LC, and JD developed the R code, performed the statistical analyses, ran the simulations, and produced the figures. JB and JD take responsibility for the accuracy of the data analysis. JB wrote the first draft of the manuscript. All authors contributed to critical revision of the report for important intellectual content and agree with the results and conclusions of this article. EB was the project manager. JP was the project supervisor. JB and JP were responsible for the final decision to submit for publication.

Declaration of interests

JP has received research support from Takeda (formerly Shire) and Aevi Genomics; and consultancy fees from Innovative Science Solutions. All other authors declare no competing interests.

Data sharing

The Adolescent Brain and Cognitive Development Study dataset, release 2.0.1, is freely available from the National Institute of Mental Health Data Archive at https://nda.nih.gov/. The R code we developed to load and process the dataset, perform the main and alternative analyses, and run the simulations is available at https://github.com/jabernanke/lancet.

Acknowledgments

JB was supported by grant number T32 MH016434-41, by the AACAP Pilot Research Award for Attention Disorders, and by the Bender-Fishbein Award in Child Psychiatry Research. JP was supported by the Edwin S Webster Foundation, Suzanne Crosby Murphy Endowment, Translating Duke Health Initiative, and by the National Institutes of Health (UH3 OD023328). We thank the Adolescent Brain and Cognitive Development study team.

References

1 Danielson ML, Bitsko RH, Ghandour RM, Holbrook JR, Kogan MD, Blumberg SJ. Prevalence of parent-reported ADHD diagnosis and associated treatment among U.S. children and adolescents, 2016. *J Clin Child Adolesc Psychol* 2018; **47:** 199–212.

2 Gaub M, Carlson CL. Gender differences in ADHD: a meta-analysis and critical review. *J Am Acad Child Adolesc Psychiatry* 1997; **36:** 1036–45.

3 Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am J Psychiatry* 2007; **164:** 942–48.

4 Thomas R, Sanders S, Doust J, Beller E, Glasziou P. Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. *Pediatrics* 2015; **135:** e994–1001.

5 Jensen CM, Steinhausen H-C. Comorbid mental disorders in children and adolescents with attention-deficit/hyperactivity disorder in a large nationwide study. *Atten Defic Hyperact Disord* 2015; **7:** 27–38.

6 Plizska SR. Treating ADHD and comorbid disorders: psychosocial and psychopharmacological interventions. New York: Guilford Press, 2009.

7 Biederman J. Attention-deficit/hyperactivity disorder: a selective overview. *Biol Psychiatry* 2005; **57:** 1215–20.

8 Tung I, Li JJ, Meza JI, et al. Patterns of comorbidity among girls with ADHD: a meta-analysis. *Pediatrics* 2016; **138:** e20160430.

9 Batty MJ, Liddle EB, Pitiot A, et al. Cortical gray matter in attention-deficit/hyperactivity disorder: a structural magnetic resonance imaging study. *J Am Acad Child Adolesc Psychiatry* 2010; **49:** 229–38.

10 Rosch KS, Crocetti D, Hirabayashi K, Denckla MB, Mostofsky SH, Mahone EM. Reduced subcortical volumes among preschool-age girls and boys with ADHD. *Psychiatry Res Neuroimaging* 2018; **271:** 67–74.

11 Ioannidis JP. Excess significance bias in the literature on brain volume abnormalities. *Arch Gen Psychiatry* 2011; **68:** 773–80.

12 Valera EM, Faraone SV, Murray KE, Seidman LJ. Meta-analysis of structural imaging findings in attention-deficit/hyperactivity disorder. *Biol Psychiatry* 2007; **61:** 1361–69.

13 Nakao T, Radua J, Rubia K, Mataix-Cols D. Gray matter volume abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication. *Am J Psychiatry* 2011; **168:** 1154–63.

14 Frodl T, Skokauskas N. Meta-analysis of structural MRI studies in children and adults with attention deficit hyperactivity disorder indicates treatment effects. *Acta Psychiatr Scand* 2012; **125:** 114–26.

15 Durham EL, Jeong HJ, Moore TM, et al. Association of gray matter volumes with general and specific dimensions of psychopathology in children. *Neuropsychopharmacology* 2021; **46:** 1333–39.

16 Hoogman M, Bralten J, Hibar DP, et al. Subcortical brain volume differences of participants with ADHD across the lifespan: an ENIGMA collaboration. *Lancet Psychiatry* 2017; **4:** 310–19.

17 Hoogman M, Muetzel R, Guimaraes JP, et al. Brain imaging of the cortex in ADHD: a coordinated analysis of large-scale clinical and population-based samples. *Am J Psychiatry* 2019; **176:** 531–42.

18 Casey BJ, Cannonier T, Conley MI, et al. The Adolescent Brain Cognitive Development (ABCD) study: imaging acquisition across 21 sites. *Dev Cogn Neurosci* 2018; **32:** 43–54.

19 Garavan H, Bartsch H, Conway K, et al. Recruiting the ABCD sample: design considerations and procedures. *Dev Cogn Neurosci* 2018; **32:** 16–22.

20 Dick AS, Lopez DA, Watts AL, et al. Meaningful associations in the adolescent brain cognitive development study. *Neuroimage* 2021; **239:** 118262.

21 Hagler DJ Jr, Hatton S, Cornejo MD, et al. Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *Neuroimage* 2019; **202:** 116091.

22 ABCD Study. ABCD protocols. https://abcdstudy.org/scientists/protocols (accessed Aug 1, 2021).

23 National Institutes of Health, National Institute on Aging. The NIH Toolbox: setting a standard for biomedical research. 2015. https://www.nia.nih.gov/research/blog/2015/10/nih-toolboxr-setting-standard-biomedical-research (accessed Aug 1, 2021).

24 Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc* 2007; **82:** 591–605.

25 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995; **57:** 289–300.

26 Aebi M, Winkler Metzke C, Steinhausen HC. Accuracy of the DSM-oriented attention problem scale of the child behavior checklist in diagnosing attention-deficit hyperactivity disorder. *J Atten Disord* 2010; **13:** 454–63.

27 Spencer AE, Plasencia N, Sun Y, et al. Screening for attention-deficit/hyperactivity disorder and comorbidities in a diverse, urban primary care setting. *Clin Pediatr (Phila)* 2018; **57:** 1442–52.

28 Sun H, Chen Y, Huang Q, et al. Psychoradiologic utility of MR imaging for diagnosis of attention deficit hyperactivity disorder: a radiomics analysis. *Radiology* 2018; **287:** 620–30.

29 Kubera KM, Schmitgen MM, Maier-Hein KH, Thomann PA, Hirjak D, Wolf RC. Differential contributions of cortical thickness and surface area to trait impulsivity in healthy young adults. *Behav Brain Res* 2018; **350:** 65–71.

30 Ellison-Wright I, Ellison-Wright Z, Bullmore E. Structural brain change in attention deficit hyperactivity disorder identified by meta-analysis. *BMC Psychiatry* 2008; **8:** 51.

31 van Erp TGM, Walton E, Hibar DP, et al. Cortical brain abnormalities in 4474 individuals with schizophrenia and 5098 control subjects via the enhancing neuro imaging genetics through meta analysis (ENIGMA) consortium. *Biol Psychiatry* 2018; **84:** 644–54.

32 Onitsuka T, Shenton ME, Salisbury DF, et al. Middle and inferior temporal gyrus gray matter volume abnormalities in chronic schizophrenia: an MRI study. *Am J Psychiatry* 2004; **161:** 1603–11.

33 Rimol LM, Nesvåg R, Hagler DJ Jr, et al. Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder. *Biol Psychiatry* 2012; **71:** 552–60.

34 Couvy-Duchesne B, Strike LT, de Zubicaray GI, et al. Lingual gyrus surface area is associated with anxiety-depression severity in young adults: a genetic clustering approach. *eNeuro* 2018; **5:** ENEURO.0153-17.2017.

35 Costa Dias TG, Iyer SP, Carpenter SD, et al. Characterizing heterogeneity in children with and without ADHD based on reward system connectivity. *Dev Cogn Neurosci* 2015; **11:** 155–74.

36 Castellanos FX, Lee PP, Sharp W, et al. Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *JAMA* 2002; **288:** 1740–48.

37 Spencer TJ, Brown A, Seidman LJ, et al. Effect of psychostimulants on brain structure and function in ADHD: a qualitative literature review of magnetic resonance imaging-based neuroimaging studies. *J Clin Psychiatry* 2013; **74:** 902–17.

38 Shaw P, Eckstrand K, Sharp W, et al. Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proc Natl Acad Sci USA* 2007; **104:** 19649–54.

39 Musser ED, Raiker JS Jr. Attention-deficit/hyperactivity disorder: an integrated developmental psychopathology and research domain criteria (RDoC) approach. *Compr Psychiatry* 2019; **90:** 65–72.

40 Nigg JT, Karalunas SL, Feczko E, Fair DA. Toward a revised nosology for attention-deficit/hyperactivity disorder heterogeneity. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2020; **5:** 726–37.