

Attention Deficit Hyperactivity Disorder Symptoms and Brain Morphology:

Examining Confounding Bias

Lorenza Dall'Aglio^{1,2*}; Hannah H. Kim^{3*}; Sander Lamballais^{4*}; Jeremy Labrecque⁵; Ryan L. Muetzel^{1†}; Henning Tiemeier^{1,3†}

¹Department of Child and Adolescent Psychiatry, Erasmus MC University Medical Center Rotterdam-Sophia Children's Hospital, Rotterdam, the Netherlands

²The Generation R Study Group, Erasmus MC University Medical Center Rotterdam, Rotterdam, the Netherlands

³Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, USA

⁴Department of Clinical Genetics, Erasmus MC University Medical Center Rotterdam, Rotterdam, the Netherlands

⁵Department of Epidemiology, Erasmus MC University Medical Center Rotterdam, Rotterdam, the Netherlands

*Co-first author

†Co-last author

Corresponding author: Henning Tiemeier, MD Ph.D.; Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, 02115 MA, USA; tiemeier@hsph.harvard.edu

ABSTRACT (199/200 words)

Background: Associations between attention-deficit/hyperactivity disorder (ADHD) and brain morphology have been reported, although with several inconsistencies. These may partly stem from confounding bias, which could distort associations and limit generalizability. We examined how associations between brain morphology and ADHD symptoms change with adjustments for potential confounders typically overlooked in the literature (aim 1), and for IQ, which is typically corrected for but plays an unclear role (aim 2).

Methods: Participants were 10-year-old children from the Adolescent Brain Cognitive Development ($N=7,961$) and Generation R ($N=2,531$) studies. Cortical area and volume were measured with MRI and ADHD symptoms with the Child Behavior Checklist. Surface-based cross-sectional analyses were run.

Results: ADHD symptoms related to widespread cortical regions when solely adjusting for demographic factors. Additional adjustments for socioeconomic and maternal behavioral confounders (aim 1) generally attenuated associations, as cluster sizes halved and effect sizes substantially reduced. Cluster sizes were further reduced when including IQ (aim 2), however, we argue that adjustments could have introduced bias (e.g., by conditioning on a collider).

Conclusions: Careful confounder selection and control can help identify more robust and specific regions of associations for ADHD symptoms, across two cohorts. We provided guidance to minimizing confounding bias in psychiatric neuroimaging.

Funding: Authors are supported by an NWO-VICI grant (NWO-ZonMW: 016.VICI.170.200 to HT) for HT, LDA, SL, and the Sophia Foundation S18-20, and Erasmus University and Erasmus MC Fellowship for RLM.

1 **Introduction**

2 Large strides have been made in the identification of neuroanatomical correlates of
3 psychiatric problems, with Attention-Deficit/Hyperactivity Disorder (ADHD) being a prominent
4 example. ADHD is the most prevalent neurodevelopmental disorder in children worldwide
5 and is characterized by atypical levels of inattention, hyperactivity, and/or impulsivity (1).
6 Structural Magnetic Resonance Imaging (sMRI) studies have highlighted that children with
7 ADHD show widespread morphological differences, such as in the basal ganglia (2),
8 subcortical areas (3), and frontal, cingulate, and temporal cortices compared to children
9 without the disorder (4,5).

10 Consistently identifying the neuroanatomical substrate of ADHD, however, remains
11 challenging. A recent meta-analysis did not find convergence across the literature on brain
12 differences in children and adolescents with ADHD (6). One possible explanation for this
13 inconsistency is the multifaceted nature of ADHD, in which children with the disorder have
14 heterogeneous presentations on several cognitive and emotional domains, which could stem
15 from distinct brain structural substrates. Other explanations regard study design. If
16 suboptimal, it may lead to biased estimates and lack of generalizability, thus potentially
17 concealing robust and replicable relations of brain morphology with ADHD. The present
18 study focuses on confounding, a common source of bias in etiological studies.

19 Confounding bias arises when a third variable affects both the determinant
20 (independent variable) and outcome (dependent variable) of interest (i.e., is a common
21 cause) (7). Confounding leads to over- or under-estimation of the true effect between
22 determinant and outcome and can even change the direction of an association. To minimize
23 confounding bias, appropriate confounder control is paramount, although it is challenging,
24 especially in observational studies like most neuroimaging studies of ADHD. Previous
25 literature and expert knowledge can guide the identification of potential confounders (8),
26 which can then be appropriately adjusted for in regression models or using methods such as
27 restriction, standardization, or propensity scores.

28 Within neuroimaging studies of ADHD, except for a few large investigations (3,9,10),
29 studies have generally matched or adjusted for a few demographic variables (e.g., age, sex)
30 and neuroimaging metrics or parameters. Of the 19 studies included in a systematic review
31 of neuroimaging studies on ADHD (11), 17 adjusted or matched for age in their analyses, 14
32 for sex, 9 for precision variables like study site, and 8 for the intelligence quotient (IQ)
33 (**Supplementary Table 1**). Further potential confounders should, however, be considered.
34 For instance, socioeconomic status (SES) is related to both higher risk for ADHD and
35 variation in cortical brain structure (12,13). Thus, it is likely a confounder. Lack of adjustment
36 for SES may have therefore concealed key relations between ADHD and brain structure.
37 Adjustment choices are dependent on the availability of large samples with data on a wide
38 variety of covariates, which has to date been limited for psychiatric neuroimaging studies.
39 Yet, this is rapidly changing with the advent of population neuroscience, which entails large-
40 scale studies with neurobiological data. This lends new opportunities for further confounder
41 adjustments to be considered in neuroimaging studies of ADHD. Conversely, previous
42 studies have adjusted for IQ, which may not be a confounder in the association between
43 ADHD symptoms and the brain, and may thus have led to further bias in the results (14).

44 In this study, we examined the association between brain structure and ADHD
45 symptoms and how the selection and control for potential confounders may affect results
46 (aim 1). Moreover, we discussed the unclear role of IQ in brain structure – ADHD
47 associations and the potential consequences of adjusting for it (aim 2). We leveraged two
48 large, population-based cohorts: the Adolescent Brain Cognitive Development (ABCD) and
49 the Generation R Studies. In line with most neuroimaging studies, we adopted a cross-
50 sectional design.

51

52 **Results**

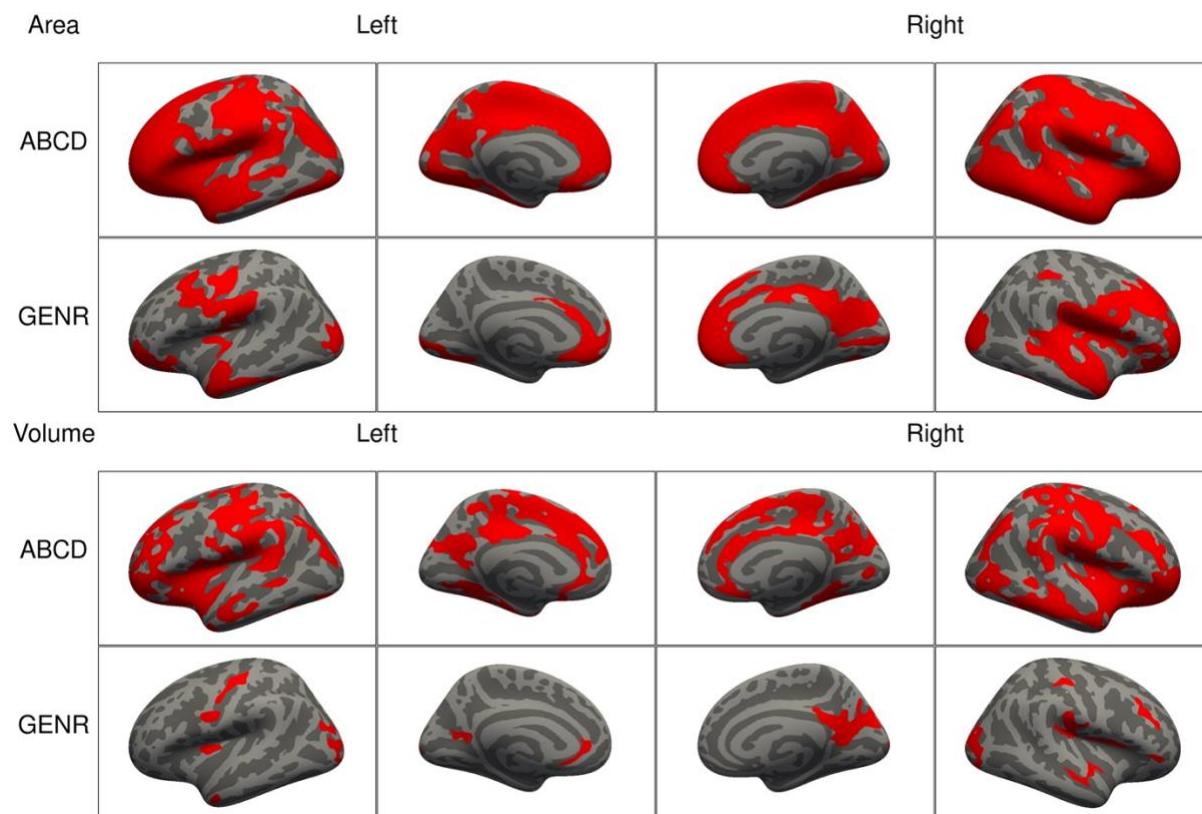
53 **Associations between ADHD symptoms and brain morphology are widespread**

54 We analyzed data from 10-year-old children from the ABCD ($N = 7,961$, multi-site)
55 and Generation R ($N = 2,531$, single-site) Studies (**Supplementary Table 2**). ADHD
56 symptoms were measured with the Child Behavioral Checklist (CBCL). T_1 -weighted images
57 were obtained with 3T scanners (15,16). Cortical surface area and volume were considered,
58 while thickness was not tested due to previously reported null findings in the Generation R
59 Study (4). We ran vertex-wise linear regression models for ADHD with cortical surface area
60 and volume. We adjusted for demographic and study characteristics which have been
61 generally considered by previous literature (**Supplementary Table 1**): age, sex, ethnicity,
62 and study site (ABCD only). We refer to this model as model 1, as further adjustments for
63 confounders are outlined in subsequent steps.

64 We found that higher ADHD symptoms were associated with less bilateral surface
65 area and cortical volume in both cohorts. As shown in **Figure 1**, associations were
66 widespread for both surface area (ABCD = $1,109.8 \text{ cm}^2$; GenR = 444.7 cm^2) and volume
67 (ABCD = 666.0 cm^2 ; GenR = 96.1 cm^2). Across both cohorts, we consistently identified
68 clusters for surface area in the lateral occipital, postcentral, rostral middle and superior
69 frontal, and superior parietal cortices. For volume, we observed overlap across cohorts in the
70 cuneus, precuneus, fusiform, inferior parietal, inferior, middle, and superior temporal,
71 isthmus of the cingulate, lateral occipital, pericalcarine, pre- and post-central, and
72 supramarginal cortices.

73

74 **Figure 1.** Significant clusters in the association of ADHD symptoms with cortical surface
75 area (top) and volume (bottom) based on the ABCD and Generation R Studies, for model 1.



76
77 *Note.* Rows represent the results for the ABCD or Generation R Studies, and the columns represent the left and right
78 hemispheres. Regions in red represent significant clusters from model 1 (adjusted for sex, age, race/ethnicity).

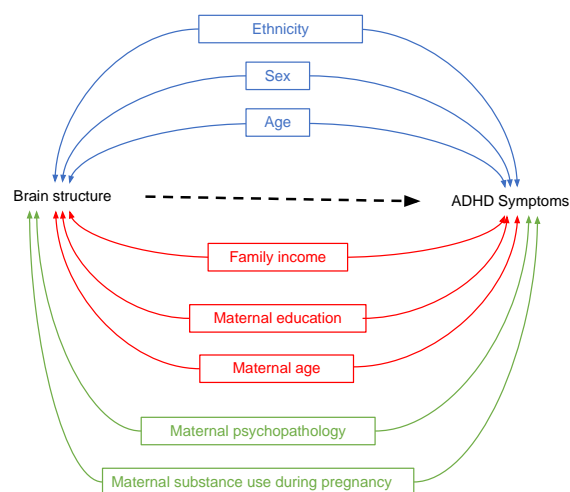
79 **Confounder selection: Socioeconomic and maternal behavioral factors**

80 Next, we considered factors that have been previously linked to ADHD and brain
81 structure in the literature, and are thus potential confounders. To illustrate this background
82 knowledge and the assumptions about relations between variables, we used Directed
83 Acyclic Graphs (DAGs), a type of causal diagram (8). These guide the identification (and
84 dismissal) of covariates that may act as confounders (**Supplementary Box 1**). Of note,
85 while assumptions may not hold, this theoretical approach is preferred to methods selecting
86 confounders based on model statistics (17). The DAGs are depicted in **Figure 2** and
87 **Supplementary Figure 1**, and the rationales for variable inclusion are explained below and
88 in the *Methods* section.

89 Based on the literature, lower SES is associated with a higher risk for ADHD (12) and
90 with variation in cortical brain structure (13). Thus, confounding by socioeconomic factors in
91 the relation between ADHD and brain morphology is likely. We therefore additionally
92 adjusted for a second set of confounders (model 2) related to SES: household income,
93 maternal education, and maternal age at childbirth.

94 Moreover, several factors concerning maternal behavior, pre- and postnatally, have
95 been associated with both ADHD and brain morphology. For instance, prenatal exposure to
96 substances is known to increase the risk of developing ADHD symptoms and has been
97 associated with variation in cerebral volume and surface area (18,19). Postnatal maternal
98 psychopathology has been linked to higher child ADHD symptoms (20) and smaller brain
99 volume in children (21). Thus, in model 3 we additionally adjusted for prenatal exposure to
100 substance use (tobacco and cannabis), and postnatal maternal psychopathology.

101 **Figure 2.** Directed Acyclic Graphs for brain structure and ADHD symptoms (simplified).



102

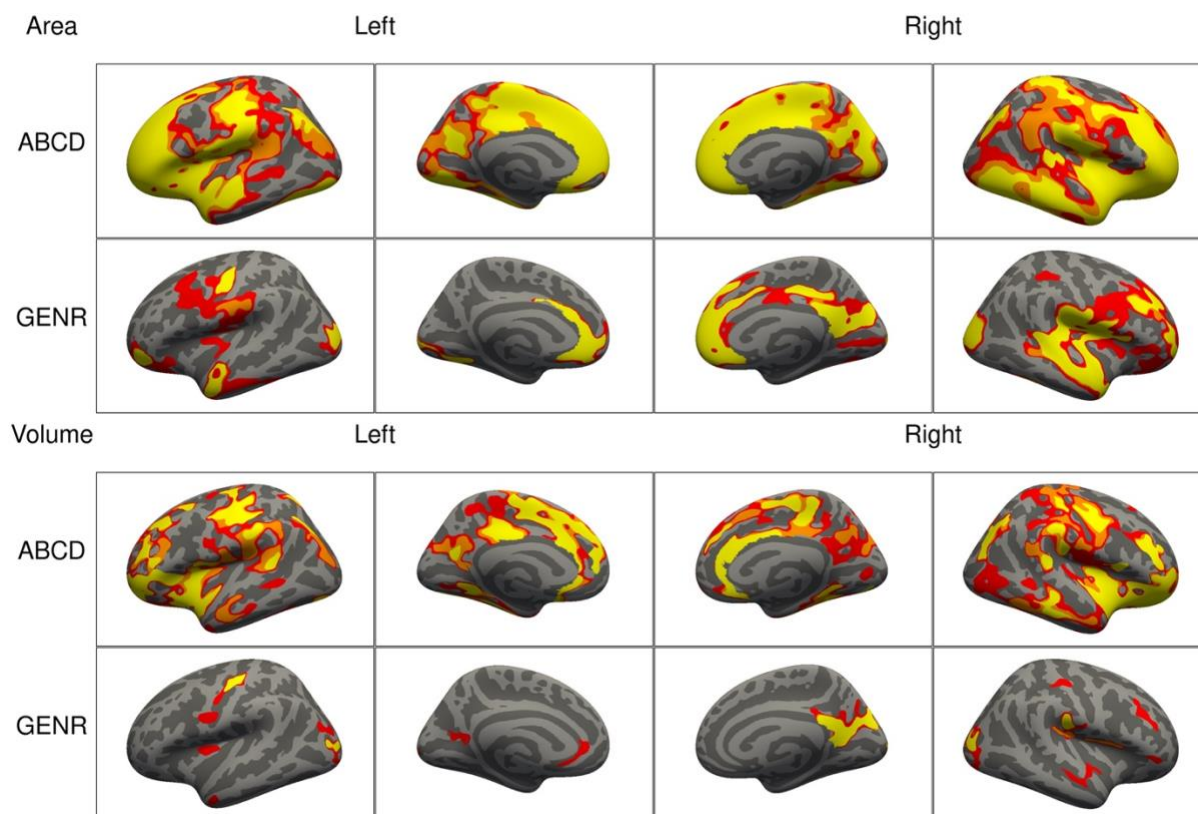
103 *Note.* DAGs illustrating potential confounders in the association between brain structure and ADHD symptoms for three
104 *sequential* models. Model 1 included demographic and study characteristics: Sex, age, ethnicity, study site (ABCD only) (in
105 blue). Model 2 additionally included socioeconomic status factors: Family income, maternal education, and maternal age at
106 childbirth (in red). Model 3 additionally incorporated postnatal maternal psychopathology and maternal substance use during
107 pregnancy (in green).

108 **Adjusting for additional confounders led to reductions in the clusters of association**

109 Adjustments for SES (model 2) led to reductions in the spatial extent of the clusters
110 for surface area and volume in both cohorts (**Figure 3**). For surface area, cluster sizes for
111 ADHD symptoms reduced from 1,109.8 cm² in model 1 to 885.4 cm² in model 2 (= - 20%) in
112 the ABCD Study, and from 444.7 cm² to 226.0 cm² (= - 49%) in the Generation R Study. In
113 the analyses of volume, clusters for ADHD symptoms were reduced from 666.0 cm² in model
114 1 to 384.6 cm² in model 2 (= - 42%) in the ABCD Study, and from 96.1 cm² to 31.6 cm² (= -
115 67%) in the Generation R Study. Adjustments for maternal substance use and
116 psychopathology (model 3) showed further cluster changes for surface area and volume in
117 both cohorts (**Figure 3**). For surface area, clusters covered 682.6 cm² in the ABCD study (= -
118 23%, compared to model 2) and 214.1 cm² (= - 5%) in the Generation R Study. Cortical
119 volume clusters related to ADHD symptoms comprised 262.5 cm² (= - 32%, compared to
120 model 2) in the ABCD Study, and 33.8 cm² (= + 7%) in the Generation R Study.

121 After adjusting for the confounders added in model 3, across both cohorts, we
122 consistently identified clusters for surface area in the cuneus, precuneus, fusiform, inferior
123 parietal, isthmus of the cingulate, pericalcarine, pre- and post-central, rostral middle and
124 superior frontal, superior temporal and supramarginal cortices. Clusters consistently
125 identified across cohorts for volume were in the inferior parietal, isthmus of the cingulate,
126 later occipital, pre- and post-central, precuneus, and supramarginal cortices.

127 **Figure 3.** Significant clusters in the association of ADHD symptoms with cortical surface
128 area (top) and volume (bottom) based on the ABCD and Generation R Studies, for models 1
129 to 3.



130

131 *Note.* Rows represent the results for the ABCD or Generation R Studies, and the columns represent the left and right
132 hemispheres. The colors denote the different models. Regions in red represent significant clusters from model 1 (sex, age,
133 race/ethnicity), orange from model 2 (model 1 + family income, maternal education, and maternal age at childbirth), and yellow
134 from model 3 (model 2 + maternal smoking, substance use during pregnancy, psychopathology).

135 **Similar results were observed for ADHD diagnosis**

136 To explore whether the results observed for associations between brain morphology and
137 ADHD symptoms applied to children with an ADHD diagnosis, we repeated the primary
138 analysis using the ADHD diagnostic data from the Kiddie Schedule for Affective Disorders
139 and Schizophrenia (KSADS) in the ABCD Study. In line with our primary results, ADHD
140 diagnosis was associated with less bilateral surface area and volume. Compared to clusters
141 for ADHD symptoms, those associated with ADHD diagnosis were smaller, but overlapping
142 (**Supplementary Figure 2**). We observed similar patterns of reduction in the spatial extent
143 of the clusters after adjusting for each set of confounders (**Supplementary Figure 3**). For
144 surface area, cluster sizes for ADHD symptoms covered 275.5 cm² in model 1 and reduced
145 to 233.0 cm² in model 2 (= - 15%), and 94.9 cm² in model 3 (= - 59%, compared to model 2).
146 For volume, cluster sizes for ADHD symptoms comprised 98.9 cm² in model 1 and reduced
147 to 70.7 cm² in model 2 (= - 29%), and 26.2 cm² in model 3 (= - 63%, compared to model 2).

148

149 **Beta coefficients generally decreased after confounder adjustments, but may also**
150 **increase**

151 Surface-based studies generally focus on the spatial extent of cortical clusters
152 associated with the phenotype, but, in this study, we also explored how confounding
153 adjustments affected the vertex-wise regression coefficients for ADHD symptoms.

154 At a vertex-wise level, adjusting for socioeconomic and maternal factors (model 3)
155 led to reductions in the beta coefficients, across the brain, for both cohorts (**Supplementary**
156 **Figure 4**). Of note, some beta coefficients also showed increases.

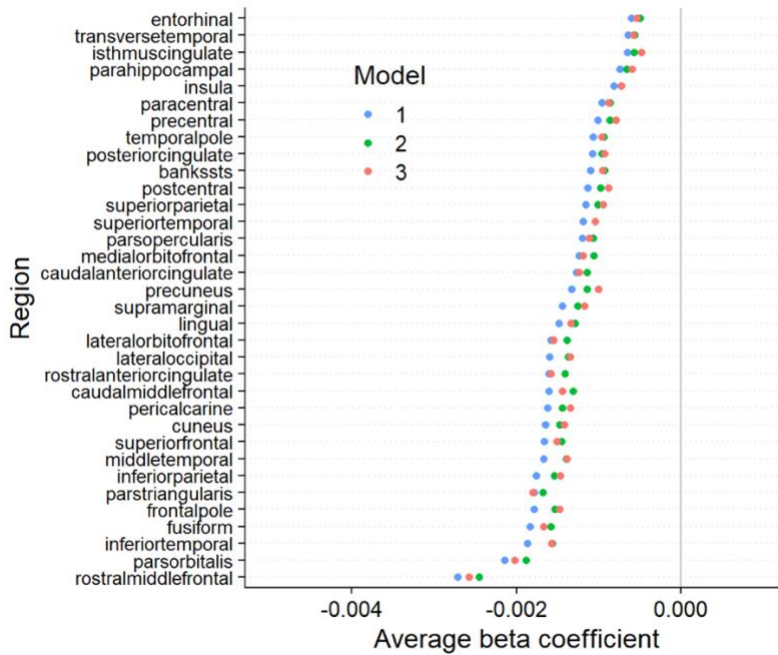
157 As confounding bias may lead to under- or over-estimation, it is not surprising to observe
158 both decreases and increases in the average beta coefficients after adjustments.

159 At an anatomical region level, where estimates of vertices within a given Desikan-
160 Killiany region were averaged, beta coefficients for surface area tended to decrease from
161 model 1 to 2 by approximately 15% (**Figure 4, Supplementary Figure 5**). Further
162 adjustments from model 2 to 3 led to decreases in the average beta coefficients of certain

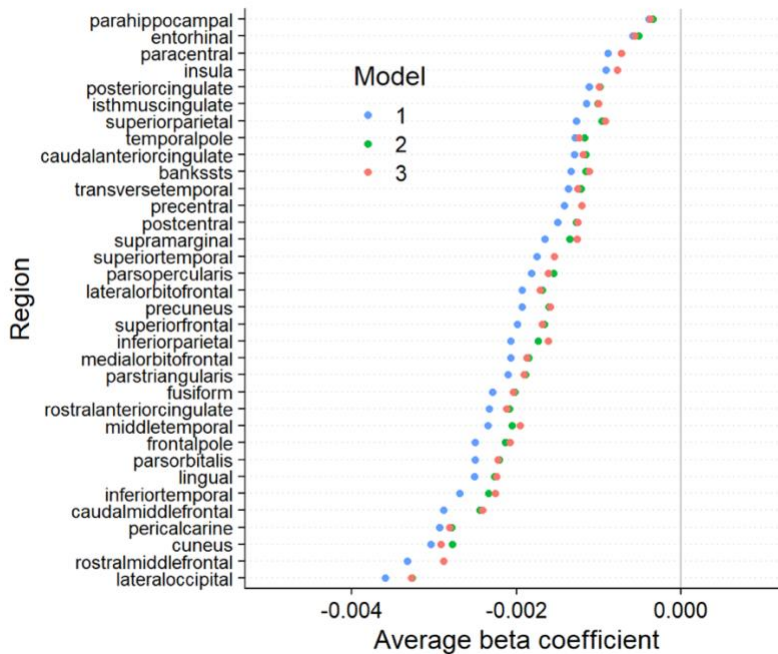
163 regions and increases in others. Similar patterns were found for volume across both studies
164 **(Supplementary Figure 6, 7)**. The average beta coefficients per region correlated
165 moderately to strongly between the ABCD and Generation R Studies for both surface area
166 ($r_{M1} = 0.83$, $r_{M2} = 0.79$, $r_{M3} = 0.77$) and volume ($r_{M1} = 0.57$, $r_{M2} = 0.56$, $r_{M3} = 0.61$)
167 **(Supplementary Figure 8)**.

168 **Figure 4.** Region-based average regression coefficients for surface area in the ABCD and
169 Generation R Studies.

ABCD



GENR



170
171 *Note.* The colors denote the different models, and the circles denote the average of all the betas within that region. The regions
172 are based on the Desikan-Killiany atlas. Results for the ABCD and Generation R Studies are respectively shown on the top and
173 bottom.

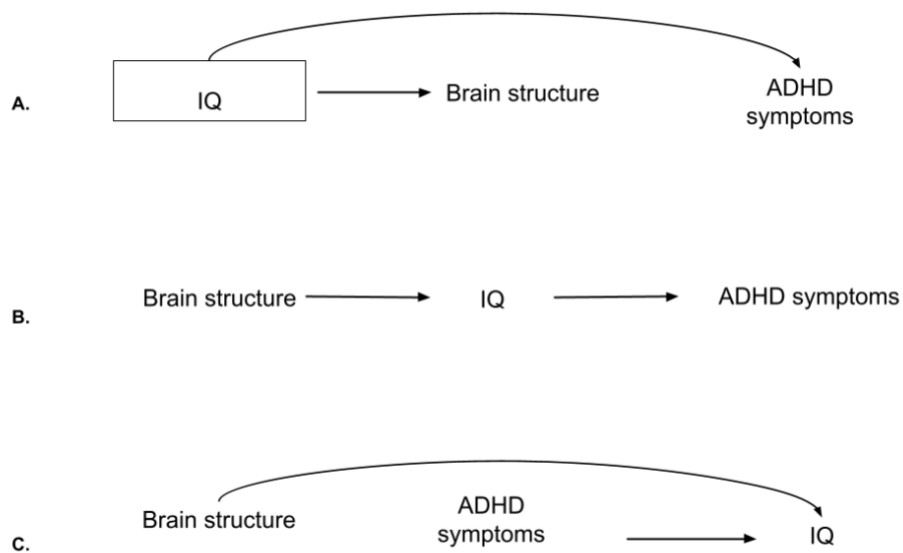
174

175 **IQ may be a confounder, mediator, or collider in neuroanatomical studies of ADHD**

176 We considered one additional scenario which included IQ, a factor that is often
177 adjusted for in previous studies (**Supplementary Table 1**). However, based on prior
178 literature, it holds an ambiguous role in structural anatomy - ADHD relations. Previous
179 studies found that children with ADHD scored lower on IQ than children without ADHD (22).
180 Differential brain structure with levels of IQ has also been shown (23). However, the
181 directions of causation between these variables remain unclear (24). IQ may therefore be a
182 confounder, collider, and/or mediator in the relation between brain structure and ADHD, as
183 depicted in the DAGs in **Figure 5** and **Supplementary Box 2**.

184 First, it could be argued that IQ is partly innate and precedes brain development and
185 ADHD, making it a confounder (**Figure 5A**). Second, IQ may lie in the pathway between
186 brain structure and ADHD and therefore act as a mediator (**Figure 5B**). It is conceivable that
187 cognitive differences, as a consequence of subtle neurodevelopmental differences (25),
188 could underlie ADHD. Adjusting for a mediator would lead to bias when estimating the total
189 association between brain structure and ADHD (26). Third, brain structure may impact
190 intelligence scores (25), and ADHD symptoms may affect IQ test performance (27) (**Figure**
191 **5C**). A variable that is independently caused by the outcome and the determinant is also
192 known as a collider, and adjusting for it leads to (collider) bias. Here, we explored the impact
193 of adjusting for IQ when examining the relation between brain morphology and ADHD
194 (model 4).

195 **Figure 5.** Directed acyclic graphs for IQ, brain structure, and ADHD symptoms



196

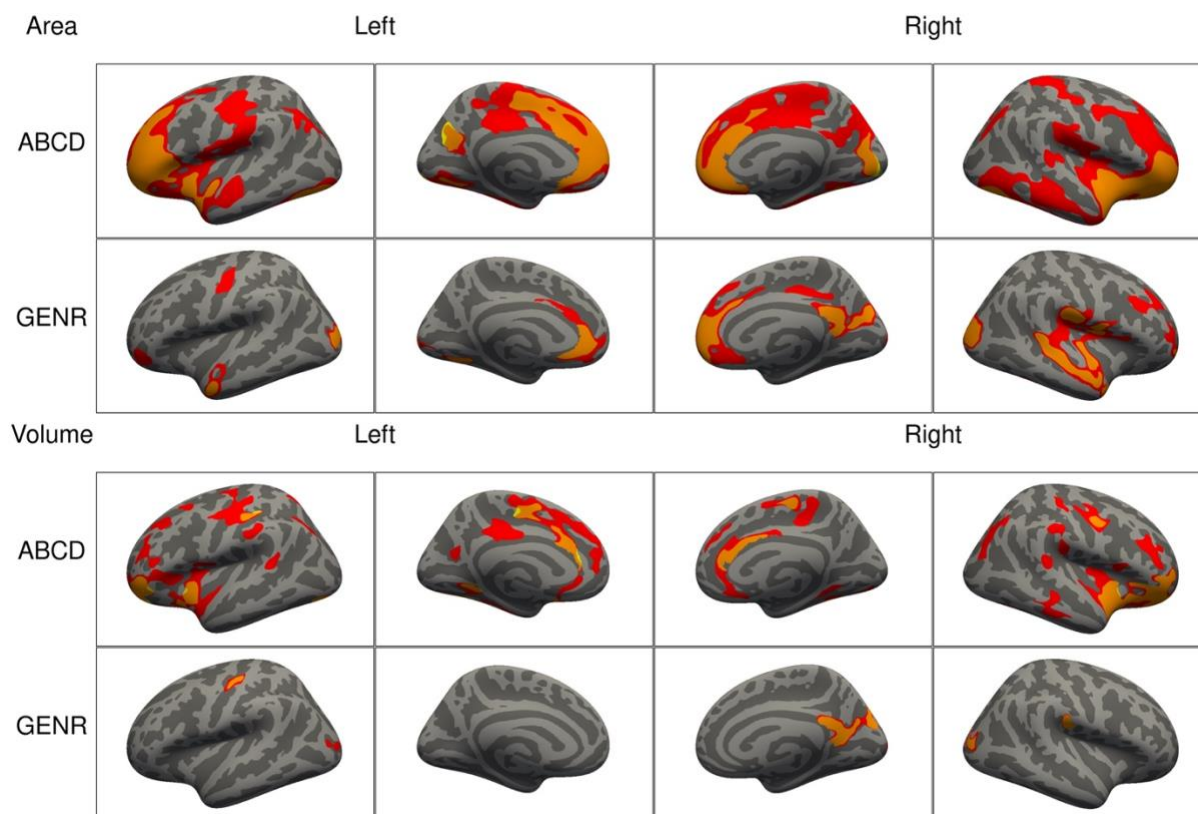
197 *Note.* **A)** DAG for IQ as a confounder. In this case, adjustments are needed as the backdoor path from brain structure to ADHD
198 symptoms through IQ is open. By adjusting (box around IQ), the path gets closed. **B)** DAG for IQ as a mediator. Adjustments
199 are not needed to estimate the total effect of brain structure on ADHD symptoms. **C)** DAG for IQ as a collider. The backdoor
200 path through IQ is already closed. Adjustments would open the path and lead to collider bias.

201 **Adjustments for IQ led to further cluster reductions**

202 After additionally adjusting for IQ, the spatial extent of the clusters associated with
203 ADHD symptoms reduced further in both cohorts (**Figure 6**). For surface area, compared to
204 model 3, clusters reduced from 682.6 cm² to 525.7 cm² (= - 23%) for the ABCD Study, and
205 from 214.1 cm² to 93.1 cm² for the Generation R Study (= - 57%). For volume, compared to
206 model 3, clusters reduced from 262.5 cm² to 164.1 cm² (= - 37%) in the ABCD Study and
207 from 33.8 cm² to 17.9 cm² (= - 47%) in the Generation R Study.

208 For both cohorts, clusters of association for surface area in model 4 were located in
209 the fusiform, inferior parietal, insula, lateral occipital, middle temporal, pericalcarine, pre- and
210 post-central, precuneus, rostral middle, and superior frontal, superior parietal and temporal,
211 and supramarginal cortices. For volume, the remaining clusters consistently identified across
212 cohorts were the isthmus of the cingulate, lateral occipital, pre- and post-central, lingual,
213 precuneus, superior parietal, and supramarginal cortices.

214 **Figure 6.** Significant clusters in the association of ADHD symptoms with cortical surface
215 area (top) and volume (bottom) based on the ABCD and Generation R Studies, after
216 additional adjustment for IQ.



217
218 *Note.* Rows represent the results for the ABCD or Generation R Studies, and the columns represent the left and right
219 hemispheres. The colors denote the different models, with red vertices being significant only in model 3, orange ones in both
220 model 3 and after adjustment for IQ, and yellow ones only after adjusting for IQ.

221

222 **Discussion**

223 By leveraging two large population-based studies and adopting a literature- and
224 DAG-informed approach to address confounding, we showed that (i) associations between
225 brain structure and ADHD symptoms, which were initially widespread, reduced when
226 adjusting for socioeconomic and maternal behavioral confounders, and that (ii) careful
227 considerations are needed when including IQ due to its unclear relation with ADHD and
228 brain morphology.

229

230 **Adjustments for confounders highlighted key regions of association, observed across** 231 **two large cohorts**

232 Widespread associations between surface area and volume with ADHD symptoms
233 were initially identified, with higher symptoms relating to smaller brain structures, in line with
234 previous research (4,28). After adjustments for potential confounders typically overlooked by
235 previous literature (socioeconomic and maternal behavioral factors), approximately half of
236 the associations remained, and considerable effect size changes were observed in both the
237 ABCD and Generation R Studies. We observed similar patterns of cluster reductions for the
238 relation of ADHD diagnosis with surface area and volume in the ABCD Study.

239 Regions that remained associated after adjustments and which were consistently
240 identified across cohorts were the precuneus, isthmus of the cingulate, supramarginal, pre-
241 and post-central, and inferior parietal cortices for both area and volume. Most of these
242 regions (e.g., supramarginal) have been previously implicated in ADHD in clinical samples
243 (29–31). However, many different brain areas have been detected in association with the
244 disorder (11), which may have hampered prior meta-analytic efforts to identify consistent
245 neuroanatomical correlates for ADHD.

246 Here, we discerned associated areas likely subject to confounding bias from areas
247 robust to socioeconomic and maternal behavioral factors, and replicable across two large
248 cohorts. Comparisons with prior findings should be made with caution due to differences in

249 study design, samples (clinical vs. population-based), and analytical methods. Importantly,
250 we highlighted the opportunity for future studies to include covariates that go beyond age
251 and sex, can help refine associations, and can be readily collected. Future studies may want
252 to consider other confounding factors, depending on their research question, design, and
253 assumed causal relations.

254

255 **Adjustments for IQ are often unnecessary when examining the relation between brain** 256 **structure and ADHD**

257 Avoiding bias from adjusting for variables that are not confounders is as important as
258 identifying sources of confounding. Adjusting for mediators or colliders of the ADHD-brain
259 structure relation would induce bias. Here, when adjusting for IQ, which plays an unclear role
260 in brain structure – ADHD associations, cluster sizes reduced considerably in both the ABCD
261 and Generation R Studies. This could indicate that IQ is a confounder, in which case
262 adjustments would be necessary, or that IQ is a mediator or collider, in which case
263 adjustments must be avoided.

264 First, based on previous literature and this study, the association between ADHD and
265 IQ is relatively weak (14) ($r_{\text{ABCD}} = -0.11$, $r_{\text{GENR}} = -0.14$), but this does not necessarily make it a
266 weak confounder as the strength of confounding is due to a variable's relation with the
267 exposure *and* outcome. Second, if brain structure and ADHD symptoms both cause
268 cognitive changes, adjusting for IQ could induce collider bias, although this is also
269 dependent on when IQ is measured relative to the exposure and outcome (8). Third, if brain
270 structure determines cognitive functioning, which in turn affects ADHD symptoms (mediation
271 by IQ), adjustments would also induce bias (26).

272 Given these scenarios, we recommend moving away from routinely adjusting for IQ
273 in ADHD neuroimaging studies, and we highlight the need to carefully consider the causal
274 model for a specific research question to determine whether IQ may confound associations.

275

276

277 **Generalization to psychiatric neuroimaging studies**

278 Our considerations on confounding likely generalize to the psychiatric neuroimaging
279 field, as several confounders considered here (e.g., SES) also relate to brain function and
280 other psychiatric disorders (32–34). Similarly, other psychiatric disorders are also
281 characterized by complex relations with IQ (35).

282 Confounding control is paramount to studies examining determinants of a phenotype,
283 like ADHD. However, even in these studies, one may be tempted to conduct correlational
284 research with limited confounding adjustments, and then speculate about biological causal
285 mechanisms (36,37). Rather, we suggest leveraging prior literature and expert knowledge to
286 identify and adjust for key confounders. This can help eliminate the influence of alternative
287 mechanisms (to the ones hypothesized) on the relation of interest (8). Charting the assumed
288 (causal) structures to identify confounders can be done through the use of tools such as
289 DAGs (8). Naturally, the plausibility of such assumptions should be evaluated. To facilitate
290 the minimization of confounding bias in psychiatric neuroimaging, we propose a workflow in
291 **Figure 7.**

Minimizing confounding bias

A workflow

When and Why

Careful confounder adjustments are paramount in studies with an etiological aim. Identify key potential confounders as early as possible and ideally before data collection. Incorporate in your study analysis techniques minimizing confounding bias (e.g. matching, adjustments)

Using models including question- and design-specific confounders minimizes incorrectly classifying effects due to confounding bias as real brain-behavior relations

Consider multiple DAGs

Where a variable may have an unclear role in the association you are examining, consider multiple DAGs with different assumptions (e.g. IQ for ADHD and the brain as confounder, collider, mediator). Of note, running multiple models will *not* inform on the role of a variable and none of the assumed models may be correct, but subject-matter knowledge and previous literature can be used to identify likely scenarios.

Interpretation

Interpret the results in light of the DAG-informed models you previously specified. When interpreting the findings, it is key to evaluate the extent to which you think your DAG model assumptions are met.

Using and building DAGs

DAGs can be useful to represent research questions and identify confounders, colliders, and mediators. DAGs can therefore guide on which variables should be adjusted for.

Information from previous literature and subject-matter knowledge can be used to guide in building a DAG. Search for information on variables which may affect both the predictor and outcome and make sure to follow DAG rules.

Consider unmeasured or residual confounding

Discuss the potential biases that may arise from variables not captured in the data (unmeasured confounding) and from measurement error in confounders (residual confounding). Consider sensitivity analyses (e.g. bounding factor, simulations) to quantify these.

294 **Limitations of the present study and suggestions for future research**

295 Despite leveraging two large samples with similar characteristics and assessments,
296 this study presents several limitations. First, there is always potential for residual
297 confounding through unmeasured confounders and misclassification of measured
298 confounders. For example, given that genetic factors influence both ADHD and brain
299 morphology and that there is a genetic correlation between ADHD risk and intracranial
300 volume (38–40), certain genetic risk variants may be unmeasured confounders. However,
301 we aimed to illustrate plausible confounding bias scenarios for ADHD and brain structure,
302 and not to provide an exhaustive list of potential confounders, which may vary depending on
303 the study population and research question. Future studies should also consider bias
304 analyses to assess the impact that residual confounding may have on the study results (41).
305 Bias analyses can help understand the minimum association strength an unmeasured
306 confounder needs to have with the determinant and outcome to fully explain away the
307 findings (42). Developments may be needed, however, for their adaptation to the
308 neuroimaging field.

309 Second, due to our cross-sectional design, deliberately chosen to correspond to most
310 neuroimaging studies, we must assume all confounders precede our determinant and
311 outcome. This is a plausible assumption for the Generation R Study as, being a prospective
312 birth cohort, we could ensure that the confounders here considered temporally preceded
313 both ADHD and neuroanatomical assessments. However, this was not possible for the
314 ABCD Study, which started sampling at child ages 9-10 years. Future research on the
315 temporal relations between potential confounders, ADHD, and brain structure will aid the
316 minimization of confounding bias when investigating the structural substrates of ADHD.

317 Third, neuroimaging parameters, such as head motion, were not considered here
318 because they are discussed elsewhere (43), and generally fall under measurement error
319 (information bias) rather than confounding bias. For instance, even when children with
320 ADHD move more in the scanner compared to controls, determining lower image quality,
321 head motion during scanning cannot cause changes in ADHD symptom levels. Future

322 research aiming to increase precision in their estimates may, however, benefit from
323 adjustments for this and other neuroimaging parameters.

324 Lastly, while we leveraged both symptom-level and diagnostic data for ADHD, this
325 was done within population-based studies. Our results cannot, therefore, be generalized to a
326 clinical population. Future research could examine the extent to which associations between
327 brain structure and ADHD change after adjustments for likely confounders in clinical
328 samples.

329 In conclusion, leveraging an empirical example from two large studies on
330 neuroanatomy and ADHD symptoms, we highlighted the opportunity for future studies to
331 consider further key confounders. These can be identified based on prior literature and
332 causal diagrams as well as be readily collected, offering a feasible venue for future research.
333 Adjusting for these potential confounders helped identify more refined cortical associations
334 with ADHD symptoms, robust to the influence of demographic and socioeconomic factors,
335 pregnancy exposures, and maternal psychopathology. We also evaluated the potential role
336 of IQ, which could be a mediator, collider, and/or confounder. While adjusting for IQ led to
337 reductions in associations, these would, however, likely not be attributable to reduced
338 confounding bias. We discussed the generalizability of these considerations on confounding
339 bias to psychiatric neuroimaging, and suggest a workflow that can be followed to minimize
340 confounding bias in future studies.

341 **Methods and Materials**

342 **Participants**

343 We analyzed data from two independent population-based cohorts: The ABCD study
344 and the Generation R Study. The ABCD study is conducted across 21 study sites in the US
345 and recruited since 2015 children aged 9 to 10 at baseline (44). The Generation R Study is
346 based in Rotterdam, the Netherlands, with data collection spanning from fetal life until early
347 adulthood, and started in 2002 (16). Details of the sampling rationale, recruitment, methods,
348 and procedures have been described elsewhere (16,44). Research protocols for the ABCD
349 study were approved by the institutional review board of the University of California, San
350 Diego, and the institutional review boards of the 21 data collection sites, while the design of
351 the Generation R study was approved by the Medical Ethics Committee of the Erasmus MC.
352 For both studies, written informed consent and assent from the primary caregiver or child
353 were obtained.

354 In this cross-sectional study, we leveraged data from the baseline assessment of the
355 ABCD study (release 2.0.1) and the 10-year assessment of the Generation R study. Both
356 waves included behavioral and neuroimaging measures. We included children with data on
357 ADHD symptoms and T₁-weighted MRI images. Participants were excluded if (i) they had
358 dental braces, (ii) incidental findings, (iii) their brain scans failed processing or quality
359 assurance procedures, or (iv) they were twins or triplets. Of note, excluding children with
360 dental braces is unlikely to determine selection bias by SES in either the ABCD or the
361 Generation R study as the former cohort covered the costs of dental braces removal for all
362 children who enrolled, while dental care is insured for all children in the Netherlands. Within
363 the Generation R study, a small set of participants were additionally excluded because they
364 had a different scan sequence. Finally, for each non-twin sibling set, one was randomly
365 included to minimize shared method variance bias. Flowcharts for participant inclusion and

366 exclusion are available in **Supplementary Figure 9**. The final samples consisted of 7,961
367 and 2,531 children from the ABCD and Generation R studies, respectively.

368

369 **Measures**

370 *ADHD symptoms*

371 Children's ADHD symptoms, reported by the primary caregiver, were measured with
372 the CBCL (school-age version) (45), an inventory widely used for parent reports of children's
373 emotional and behavioral problems. The attention problem syndrome scale (20 items)
374 measures inattention, hyperactivity, and impulsivity and has been previously shown to have
375 clinical utility and to discriminate between ADHD cases and controls (46). Attention problems
376 were analyzed on a discrete scale (range 0-19). For the ABCD study, we repeated the
377 analysis using present ADHD diagnosis from a parent-reported and computerized version of
378 the KSADS-5. This is a dimensional and categorical assessment used to diagnose current
379 and past psychiatric disorders according to the Diagnostic and Statistical Manual of Mental
380 Disorders (Fifth Edition) (47,48).

381

382 *Image acquisition*

383 T₁-weighted data were obtained on multiple 3T scanners in the ABCD study
384 (Siemens Prisma, General Electric (GE) 750 and Philips) and one scanner in the Generation
385 R study (GE MR750w). Standard adult-sized coils were used for the ABCD study and an
386 eight-channel receive-only head coil for the Generation R study. To acquire T₁-weighted
387 structural images, the ABCD study used an inversion prepared RF-spoiled gradient echo
388 scan with prospective motion correction while the Generation R study used an inversion
389 recovery fast spoiled gradient recalled sequence (GE option = BRAVO, TR = 8.77 ms, TE =
390 3.4 ms, TI = 600 ms, flip angle = 10°, matrix size = 220 × 220, field of view = 220 mm × 220
391 mm, slice thickness = 1 mm, number of slices = 230, ARC acceleration factor = 2). More
392 details can be found elsewhere (15,49,50). Of note, in the ABCD study, a technical mistake

393 occurred at one collection site, causing the hemisphere data to be flipped. This was fixed
394 before processing.

395

396 *Image processing*

397 FreeSurfer (version 6.0.0) was used for image processing, which involved (i) removal
398 of non-brain tissue, (ii) correction of voxel intensities for B₁ field inhomogeneities, (iii) tissue
399 segmentation, and (iv) cortical surface-based reconstruction. Cortical surface maps were
400 smoothed with a full width of a half-maximum Gaussian kernel of 10 mm. Within the ABCD
401 Study, quality assessment was based on the quality control and recommended inclusion
402 criteria for structural data from the ABCD team (49). Within the Generation R Study, quality
403 assurance was manually performed by visually inspecting all images by trained raters, as
404 previously described in the literature (51). Poor quality reconstructions were excluded.

405

406 *Covariate assessment*

407 *The ABCD study.* All data were collected at baseline (child age 9 to 10-years). Age
408 and sex were recorded at intake. Child race/ethnicity was reported by the primary caregiver
409 and was categorized as White, Black, Hispanic, Asian, Other by the ABCD team. Household
410 combined net income (<\$50,000, >=\$50,000 & < \$100,000, >=\$100,000) and highest
411 maternal education (<high school, high school diploma/GED, some college, bachelor
412 degree, postgraduate degree) were self-reported by the primary caregiver in the Parent
413 Demographics Survey. Maternal age at childbirth was measured in the Developmental
414 History Questionnaire. Tobacco and cannabis use during pregnancy were retrospectively
415 reported by the mother (yes, no, I do not know) in the Developmental History Questionnaire.
416 Caregiver psychopathology was obtained from the Total Problems Adult Self Report
417 Syndrome Scale. The Wechsler Intelligence Scale for Children-6 Matrix Reasoning total
418 scaled score was used as a proxy for IQ.

419 *The Generation R study.* Age and sex were measured based on medical records
420 obtained at birth. Child ethnicity (western, non-western) was assessed based on the parents'

421 birth country, in line with the Statistics Netherlands bureau. Maternal age at childbirth was
422 prospectively measured. Family income and highest maternal education were obtained
423 through prospective self-reports by the mother and/or father at child age 5 years. Maternal
424 education was coded into low (no/primary education), intermediate (secondary school,
425 vocational training), and high (Bachelor's degree/University). Household net monthly income
426 was classified as low (< 2000 euros), middle (2,000-3,200 euros), and high (> 3,200 euros).
427 Maternal postnatal psychopathology, measured at child age six months, was prospectively
428 reported by the mother based on the Brief Symptom Inventory questionnaire global severity
429 index. Mothers prospectively reported smoking (never used, used) and cannabis use during
430 pregnancy (no use vs. use during pregnancy). Non-verbal child IQ was measured at child
431 age five years, based on the Snijders-Oomen Niet-Verbale Intelligentie Test (52), a validated
432 Dutch non-verbal intelligence test.

433

434 *Covariate selection*

435 Similar covariates were grouped into confounding sets to minimize the number of
436 tested models while including relevant confounders. Factors included in model 1 related to
437 demographic and study characteristics (age, sex, ethnicity, and study site). Age and sex
438 were selected as these have been previously adjusted for in previous neuroimaging studies
439 of ADHD (**Supplementary Table 1**). Ethnicity was used as a proxy for differential health risk
440 exposure among people of different ethnic groups. The study site was incorporated to
441 account for location and scanner differences in the ABCD study.

442 Further potential confounders were selected based on previous literature and with
443 the aid of DAGs, as described in the *Results* section. In model 2, variables indicating
444 socioeconomic factors were included (maternal education, household income, maternal age
445 at childbirth). Household income and maternal education are generally considered to
446 measure childhood SES in health research (53). Maternal age at childbirth can additionally
447 inform on the SES of the child by capturing part of the variance unexplained by income and
448 education (e.g., younger mothers facing higher occupational challenges, highly educated

449 mothers delaying childbirth (54)). In model 3, maternal factors from the prenatal and
450 postnatal period were grouped (tobacco and cannabis use during pregnancy and maternal
451 psychopathology) to measure early life exposures which may impact a child's brain and
452 psychiatric development.

453

454 **Statistical Analyses**

455 The R statistical software (version 3.4.3) was used for all analyses. Missing data on
456 covariates were imputed with chained equations using the *mice* R package (55). Linear
457 vertex-wise analyses were performed with the *QDECOR* R package (56), with surface
458 area/volume and ADHD symptoms as variables of interest. Thickness was not examined due
459 to previously reported null findings in the Generation R study (4), and because it adds little
460 information on top of surface area and volume. Correction for multiple testing was applied by
461 using cluster-wise corrections based on Monte Carlo simulations with a cluster forming
462 threshold of 0.001, which yields false-positive rates similar to full permutation testing (57). A
463 Bonferroni correction was applied to adjust for analyzing both hemispheres separately (i.e., p
464 < 0.025 cluster-wise).

465 Our analyses for aim 1 involved three vertex-wise linear regression models, which
466 progressively expanded to adjust for additional confounding factors. The first model focused
467 on demographic covariates, the second on socioeconomic ones, and the third on maternal
468 behavioral variables related to psychopathology and pregnancy exposures. These models
469 were run for ADHD symptoms (in both the ABCD and Generation R studies) and ADHD
470 diagnosis (in the ABCD study only, as sensitivity analysis). One additional model, building
471 upon model 3, was run to address aim 2, to illustrate the consequences of adjusting for IQ.
472 Given that IQ and ADHD were weakly correlated ($r_{ABCD} = -0.11$, $r_{GENR} = -0.14$),
473 multicollinearity was not expected.

ACKNOWLEDGEMENTS

The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. All supporters are mentioned at <https://abcdstudy.org/federal-partners.html>. Participating sites and study investigators are shown at https://abcdstudy.org/consortium_members/. While ABCD investigators designed, implemented the study, and/or provided data, they did not participate in this manuscript. This work reflects the authors' views, and may not reflect those of the NIH or ABCD investigators. The Generation R Study is supported by Erasmus MC, Erasmus University Rotterdam, the Rotterdam Homecare Foundation, the Municipal Health Service Rotterdam area, the Stichting Trombosedienst & Artsenlaboratorium Rijnmond, the Netherlands Organization for Health Research and Development (ZonMw), and the Ministry of Health, Welfare and Sport. Neuroimaging data acquisition was funded by the European Community's 7th Framework Program (FP7/2008-2013, 212652, Nutrimenthe). Netherlands Organization for Scientific Research (Exacte Wetenschappen) and SURFsara (Cartesius Compute Cluster, www.surfsara.nl) supported the Supercomputing resources. Authors are supported by an NWO-VICI grant (NWO-ZonMW: 016.VICI.170.200 to HT) for HT, LDA, SL, and the Sophia Foundation S18-20, and Erasmus University and Erasmus MC Fellowship for RLM. We thank the participants, general practitioners, hospitals, midwives, and pharmacies in Rotterdam who contributed to the study.

DISCLOSURES

We declare no conflicts of interest.

REFERENCES

1. American Psychiatric Association (2013): *Diagnostic and Statistical Manual of Mental Disorders (5th Ed.)*.
2. Nakao T, Radua J, Rubia K, Mataix-Cols D (2011): Gray Matter Volume Abnormalities in ADHD: Voxel-Based Meta-Analysis Exploring the Effects of Age and Stimulant Medication. *Am J Psychiatry* 168: 1154–63.
3. Hoogman M, Bralten J, Hibar DP, Mennes M, Zwiers MP, Schweren LSJ, *et al.* (2017): Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry* 4. [https://doi.org/10.1016/S2215-0366\(17\)30049-4](https://doi.org/10.1016/S2215-0366(17)30049-4)
4. Hoogman M, Muetzel R, Guimaraes JP, Shumskaya E, Mennes M, Zwiers MP, *et al.* (2019): Brain Imaging of the Cortex in ADHD: A Coordinated Analysis of Large-Scale Clinical and Population-Based Samples. *Am J Psychiatry* 176: 531–542.
5. Shaw P, Malek M, Watson B, Greenstein D, de Rossi P, Sharp W (2018): Trajectories of Cerebral Cortical Development in Childhood and Adolescence and Adult Attention-Deficit/Hyperactivity Disorder. *Biol Psychiatry* 74: 599–606.
6. Samea F, Soluki S, Nejati V, Zarei M, Cortese S, Eickhoff SB, *et al.* (2019): Brain alterations in children/adolescents with ADHD revisited: A neuroimaging meta-analysis of 96 structural and functional studies. *Neurosci Biobehav Rev* 100: 1–8.
7. VanderWeele TJ (2019): Principles of confounder selection. *Eur J Epidemiol* 34: 211–219.
8. Hernan MA, Robins JM (2020): *Causal Inference What If?* CRC press Taylor & Francis Group.

9. Mous S, Muetzel R, El Marroun H, Polderman T, Lugt A, Jaddoe V, *et al.* (2014): Cortical thickness and inattention/hyperactivity symptoms in young children: A population-based study. *Psychol Med* 44: 1–11.
10. Bernanke J, Luna A, Chang L, Bruno E, Dworkin J, Posner J (2022): 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. *Lancet Psychiatry* 0. [https://doi.org/10.1016/S2215-0366\(21\)00505-8](https://doi.org/10.1016/S2215-0366(21)00505-8)
11. Saad JF, Griffiths KR, Korgaonkar MS (2020): A Systematic Review of Imaging Studies in the Combined and Inattentive Subtypes of Attention Deficit Hyperactivity Disorder. *Front Integr Neurosci* 14: 31.
12. Russell AE, Ford T, Williams R, Russell G (2016): The Association Between Socioeconomic Disadvantage and Attention Deficit/Hyperactivity Disorder (ADHD): A Systematic Review. *Child Psychiatry Hum Dev* 47: 440–458.
13. Noble KG, Houston SM, Brito NH, Bartsch H, Kan E, Kuperman JM, *et al.* (2015): Family income, parental education and brain structure in children and adolescents. *Nat Neurosci* 18: 773–778.
14. Dennis M, Francis DJ, Cirino PT, Schachar R, Barnes MA, Fletcher JM (2009): Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *J Int Neuropsychol Soc JINS* 15: 331–343.
15. Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, *et al.* (2018): The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci* 32: 43–54.

16. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IJzendoorn MH, *et al.* (2016): The Generation R Study: design and cohort update 2017. *Eur J Epidemiol* 31: 1243–1264.
17. Lee PH (2014): Should we adjust for a confounder if empirical and theoretical criteria yield contradictory results? A simulation study. *Sci Rep* 4: 6085.
18. Eilertsen EM, Gjerde LC, Reichborn-Kjennerud T, Ørstavik RE, Knudsen GP, Stoltenberg C, *et al.* (2017): Maternal alcohol use during pregnancy and offspring attention-deficit hyperactivity disorder (ADHD): a prospective sibling control study. *Int J Epidemiol* 46: 1633–1640.
19. Lees B, Mewton L, Jacobus J, Valadez EA, Stapinski LA, Teesson M, *et al.* (2020): Association of Prenatal Alcohol Exposure With Psychological, Behavioral, and Neurodevelopmental Outcomes in Children From the Adolescent Brain Cognitive Development Study. *Am J Psychiatry* 177: 1060–1072.
20. Clavarino AM, Mamun AA, O’Callaghan M, Aird R, Bor W, O’Callaghan F, *et al.* (2010): Maternal Anxiety and Attention Problems in Children at 5 and 14 Years [no. 6]. *J Atten Disord* 13: 658–667.
21. Zou R, Tiemeier H, van der Ende J, Verhulst FC, Muetzel RL, White T, *et al.* (2019): Exposure to Maternal Depressive Symptoms in Fetal Life or Childhood and Offspring Brain Development: A Population-Based Imaging Study. *Am J Psychiatry* 176: 702–710.
22. Bridgett DJ, Walker ME (2006): Intellectual functioning in adults with ADHD: A meta-analytic examination of full scale IQ differences between adults with and without ADHD. *Psychol Assess* 18: 1–14.

23. McDaniel MA (2005): Big-brained people are smarter: A meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence* 33: 337–346.
24. Gallo EF, Posner J (2016): Moving towards causality in attention-deficit hyperactivity disorder: overview of neural and genetic mechanisms. *Lancet Psychiatry* 3: 555–567.
25. Lee JJ, McGue M, Iacono WG, Michael AM, Chabris CF (2019): The causal influence of brain size on human intelligence: Evidence from within-family phenotypic associations and GWAS modeling. *Intelligence* 75: 48–58.
26. VanderWeele TJ (2016): Mediation Analysis: A Practitioner's Guide. *Annu Rev Public Health* 37: 17–32.
27. Jepsen JRM, Fagerlund B, Mortensen EL (2009): Do Attention Deficits Influence IQ Assessment in Children and Adolescents With ADHD? *J Atten Disord* 12: 551–562.
28. Gehricke J-G, Kruggel F, Thampipop T, Alejo SD, Tatos E, Fallon J, Muftuler LT (2017): The brain anatomy of attention-deficit/hyperactivity disorder in young adults – a magnetic resonance imaging study. *PLOS ONE* 12: e0175433.
29. Saad JF, Griffiths KR, Kohn MR, Clarke S, Williams LM, Korgaonkar MS (2017): Regional brain network organization distinguishes the combined and inattentive subtypes of Attention Deficit Hyperactivity Disorder. *NeuroImage Clin* 15: 383–390.
30. Lei D, Ma J, Du X, Shen G, Jin X, Gong Q (2014): Microstructural Abnormalities in the Combined and Inattentive Subtypes of Attention Deficit Hyperactivity Disorder: a Diffusion Tensor Imaging Study. *Sci Rep* 4: 6875.
31. Solanto MV, Pope-Boyd SA, Tryon WW, Stepak B (2009): Social Functioning in Predominantly Inattentive and Combined Subtypes of Children With ADHD. *J Atten Disord* 13: 27–35.

32. Biazoli CE, Salum GA, Gadelha A, Rebello K, Moura LM, Pan PM, *et al.* (2020): Socioeconomic status in children is associated with spontaneous activity in right superior temporal gyrus. *Brain Imaging Behav* 14: 961–970.
33. Kivimäki M, Batty GD, Pentti J, Shipley MJ, Sipilä PN, Nyberg ST, *et al.* (2020): Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. *Lancet Public Health* 5: e140–e149.
34. Apter G, Bobin A, Genet M-C, Gratier M, Devouche E (2017): Update on Mental Health of Infants and Children of Parents Affected With Mental Health Issues. *Curr Psychiatry Rep* 19: 72.
35. Der G, Batty GD, Deary IJ (2009): The association between IQ in adolescence and a range of health outcomes at 40 in the 1979 US National Longitudinal Study of Youth. *Intelligence* 37: 573–580.
36. Hernán MA (2018): The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data. *Am J Public Health* 108: 616–619.
37. Grosz MP, Rohrer JM, Thoemmes F (2020): The Taboo Against Explicit Causal Inference in Nonexperimental Psychology. *Perspect Psychol Sci* 15: 1243–1255.
38. Jansen AG, Mous SE, White T, Posthuma D, Polderman TJC (2015): What Twin Studies Tell Us About the Heritability of Brain Development, Morphology, and Function: A Review [no. 1]. *Neuropsychol Rev* 25: 27–46.
39. Klein M, Walters RK, Demontis D, Stein JL, Hibar DP, Adams HH, *et al.* (2019): Genetic Markers of ADHD-Related Variations in Intracranial Volume. *Am J Psychiatry* 176: 228–238.

40. Klein M, Onnink M, van Donkelaar M, Wolfers T, Harich B, Shi Y, *et al.* (2017): Brain imaging genetics in ADHD and beyond – Mapping pathways from gene to disorder at different levels of complexity. *Neurosci Biobehav Rev* 80: 115–155.
41. Lash T, Fox M, MacLehose R (2021): *Applying Quantitative Bias Analysis to Epidemiologic Data*. Springer International Publishing. Retrieved February 16, 2022, from <https://link.springer.com/book/9783030826727>
42. VanderWeele TJ, Ding P (2017): Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med* 167: 268–274.
43. Alfaro-Almagro F, McCarthy P, Afyouni S, Andersson JLR, Bastiani M, Miller KL, *et al.* (2021): Confound modelling in UK Biobank brain imaging. *NeuroImage* 224: 117002.
44. Garavan H, Bartsch H, Conway K, Decastro A, Goldstein RZ, Heeringa S, *et al.* (2018): Recruiting the ABCD sample: Design considerations and procedures. *Dev Cogn Neurosci* 32: 16–22.
45. Achenbach, Rescorla (2001): *Manual for the ASEBA School-Age Forms and Profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.
46. Eiraldi RB, Power TJ, Karustis JL, Goldstein SG (2000): Assessing ADHD and Comorbid Disorders in Children: The Child Behavior Checklist and the Devereux Scales of Mental Disorders. *J Clin Child Psychol* 29: 3–16.
47. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, *et al.* (1997): Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 36: 980–988.

48. Kobak KA (2020): *Computerized Screening for Comorbidity in Adolescents With Substance or Psychiatric Disorders*. *clinicaltrials.gov*. Retrieved January 13, 2022, from <https://clinicaltrials.gov/ct2/show/NCT01866956>
49. Hagler DJ, Hatton SeanN, Cornejo MD, Makowski C, Fair DA, Dick AS, *et al.* (2019): Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *NeuroImage* 202: 116091.
50. White T, Muetzel RL, El Marroun H, Blanken LME, Jansen P, Bolhuis K, *et al.* (2018): Paediatric population neuroimaging and the Generation R Study: the second wave. *Eur J Epidemiol* 33: 99–125.
51. Muetzel RL, Mulder RH, Lamballais S, Cortes Hidalgo AP, Jansen P, Güroğlu B, *et al.* (2019): Frequent Bullying Involvement and Brain Morphology in Children. *Front Psychiatry* 10. <https://doi.org/10.3389/fpsy.2019.00696>
52. Tellegen PJ, Laros JA (1998): Snijders-Oomen Niet-verbale Intelligentietest SON-R 2½ - 7. Retrieved June 28, 2021, from <https://research.rug.nl/en/publications/snijders-oomen-niet-verbale-intelligentietest-son-r-2-7>
53. Children, Youth, Families and Socioeconomic Status (n.d.): <https://www.apa.org>. Retrieved November 22, 2021, from <https://www.apa.org/pi/ses/resources/publications/children-families>
54. Heck KE, Schoendorf KC, Ventura SJ, Kiely JL (1997): Delayed Childbearing by Education Level in the United States, 1969–1994. *Matern Child Health J* 1: 81–88.
55. Buuren S van, Groothuis-Oudshoorn K (2011): mice: Multivariate Imputation by Chained Equations in R [no. 1]. *J Stat Softw* 45: 1–67.
56. Lamballais S, Muetzel RL (2021): QDECR: A Flexible, Extensible Vertex-Wise Analysis Framework in R. *Front Neuroinformatics* 15. <https://doi.org/10.3389/fninf.2021.561689>

57. Greve DN, Fischl B (2018): False Positive Rates in Surface-based Anatomical Analysis.

NeuroImage 171: 6–14.

STROBE Flowchart

Flowcharts of participant inclusion and exclusion for ABCD (panel A) and Generation R

(panel B).

