A Genome-Wide Association Meta-Analysis of Attention-Deficit/Hyperactivity Disorder Symptoms in Population-Based Pediatric Cohorts

Christel M. Middeldorp, MD, PhD, Anke R. Hammerschlag, MSc, Klaasjan G. Ouwens, MSc, Maria M. Groen-Blokhuis, MD, PhD, Beate St. Pourcain, PhD, Corina U. Greven, PhD, Irene Pappa, MD, PhD, Carla M.T. Tiesler, PhD, Wei Ang, PhD, Ilya M. Nolte, PhD, Natalia Vilor-Tejedor, MSc, Jonas Bacelis, MSc, Jane L. Ebejer, PhD, Huiying Zhao, PhD, Gareth E. Davies, PhD, Erik A. Ehli, PhD, David M. Evans, PhD, Iryna O. Fedko, MSc, Mónica Guxens, PhD, Jouke-Jan Hottenga, PhD, James J. Hudziak, MD, Astanand Jugessur, PhD, John P. Kemp, PhD, Eva Krapohl, MSc, Nicholas G. Martin, PhD, Mario Murcia, MSc, Ronny Myhre, PhD, Johan Ormel, PhD, Susan M. Ring, PhD, Marie Standl, PhD, Evie Stergiakouli, PhD, Camilla Stoltenberg, MD, PhD, Elisabeth Thiering, PhD, Nicholas J. Timpson, PhD, Maciej Trzaskowski, PhD, Peter J. van der Most, PhD, Carol Wang, BSc, EARly Genetics and Lifecourse Epidemiology (EAGLE) Consortium, Psychiatric Genomics Consortium ADHD Working Group, Dale R. Nyholt, PhD, Sarah E. Medland, PhD, Benjamin Neale, PhD, Bo Jacobsson, MD, PhD, Jordi Sunyer, PhD, Catharina A. Hartman, PhD, Andrew J.O. Whitehouse, PhD, Craig E. Pennell, MBBS, PhD, Joachim Heinrich, PhD, Robert Plomin, PhD, George Davey Smith, PhD, Henning Tiemeier, MD, PhD, Danielle Posthuma, PhD, Dorret I. Boomsma, PhD

Objective: The aims of this study were to elucidate the influence of common genetic variants on childhood attention-deficit/hyperactivity disorder (ADHD) symptoms, to identify genetic variants that explain its high heritability, and to investigate the genetic overlap of ADHD symptom scores with ADHD diagnosis.

Method: Within the EARly Genetics and Lifecourse Epidemiology (EAGLE) consortium, genome-wide single nucleotide polymorphisms (SNPs) and ADHD symptom scores were available for 17,666 children (<13 years of age) from nine population-based cohorts. SNP-based heritability was estimated in data from the three largest cohorts. Meta-analysis based on genome-wide association (GWA) analyses with SNPs was followed by gene-based association tests, and the overlap in results with a meta-analysis in the Psychiatric Genomics Consortium (PGC) case-control ADHD study was investigated.

Results: SNP-based heritability ranged from 5% to 34%, indicating that variation in common genetic variants influences ADHD symptom scores. The meta-analysis did not detect genome-wide significant SNPs, but three genes, lying close to each other with SNPs in high linkage disequilibrium (LD), showed a gene-wide significant association (p values between 1.46 × 10⁻⁶ and 2.66 × 10⁻⁶). One gene, WASL, is involved in neuronal development. Both SNP- and gene-based analyses indicated overlap with the PGC meta-analysis results with the genetic correlation estimated at 0.96.

Conclusion: The SNP-based heritability for ADHD symptom scores indicates a polygenic architecture, and genes involved in neurite outgrowth are possibly involved. Continuous and dichotomous measures of ADHD appear to assess a genetically common phenotype. A next step is to combine data from population-based and case-control cohorts in genetic association studies to increase sample size and to improve statistical power for identifying genetic variants.

Key words: GWA, SNP heritability, attention problems, ADHD symptoms, meta-analysis


Attention-deficit/hyperactivity disorder (ADHD) is a common psychiatric condition in childhood, with a prevalence of around 5% across countries worldwide. As an objective diagnostic test is lacking, diagnoses are based on the occurrence of age-inappropriate impulsive, hyperactive, and inattentive behaviors that occur in multiple settings and cause significant impairment. It is well established that genetic factors explain a large part of the individual differences in the vulnerability for ADHD. The heritability of childhood ADHD and related traits, such as continuous measures of attention problems and hyperactivity, has been estimated at around...
75%. Consequently, several studies now aim to identify genetic variants for ADHD. Ten candidate genes show replicated evidence for association, according to a recent review. In genome-wide association (GWA) studies, the test of the effects of single genetic variants has not yet yielded genome-wide significant hits, but a gene-enrichment analysis, including single nucleotide polymorphisms (SNPs) showing genome-wide suggestive signals, pointed to several biological pathways involved in neural processes, such as neurodevelopment. Additional evidence for the role of common SNPs (SNPs with a frequency above 5%) in ADHD comes from polygenic analyses in which the joint effect of a large number of SNPs or all SNPs is estimated. All studies but one found that the SNPs explained a significant proportion of the variance, suggesting that associated genetic variants are likely to be detected in larger GWA meta-analyses. This is confirmed by the results from the latest meta-analysis of the Psychiatric Genomics Consortium (PGC) ADHD subgroup, which yielded several genome-wide significant hits (D. Demontis for the PGC ADHD subgroup: presentation 23rd World Congress of Psychiatric Genetics, October 2015, Toronto, CA).

Another recommendation for future gene-finding studies has been to incorporate dimensional approaches of ADHD, such as continuous measures of ADHD symptoms. This is supported by polygenic risk score analyses showing that individuals’ polygenic risk scores based on the effects of SNPs in a GWA analysis in ADHD case-control studies significantly predicted continuous ADHD symptom scores and vice versa. Other studies have suggested that a diagnosis of ADHD can be regarded as the extreme end of a continuous distribution of inattentive and hyperactive behaviors, and twin studies also showed a substantial overlap between the genetic factors for a clinical diagnosis of ADHD and continuous measures of ADHD symptoms in the general population.

Many population-based pediatric cohorts have collected genome-wide SNP data and continuous ADHD symptom scores, providing an underused opportunity for gene-finding studies for ADHD. Case-control studies benefit from oversampling the high-scoring end of the distribution, but analyzing the full information on symptom severity in the population can also be a powerful approach, especially for a relatively common disorder such as ADHD. The current report describes the first GWA meta-analysis of continuous measures of ADHD in 17,666 children from nine population-based cohorts. To investigate the polygenic nature of the phenotype, we first estimated the variance in attention problems that was explained by all SNPs. Next, SNP and gene-based association analyses were performed. We specifically looked at whether the 10 candidate genes identified in the recent review showed evidence of association. Finally, to examine the overlap in genetic influences on ADHD diagnosis and ADHD symptom scores, we investigated the concordance in the effects of SNPs and genes as found in the current meta-analysis with the results of a meta-analysis based on case-control ADHD GWA studies, and estimated the genetic correlation.

### Table 1: Description of Cohorts and Attention-Deficit/Hyperactivity Disorder (ADHD) Instruments Included in Meta-Analysis

<table>
<thead>
<tr>
<th>Cohort</th>
<th>N</th>
<th>Phenotype Instrument</th>
<th>Rater</th>
<th>Age, y, mean (SD)</th>
<th>Sum Score mean (SD)</th>
<th>Web Site</th>
<th>Reference Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALSPAC</td>
<td>5,757</td>
<td>SDQ</td>
<td>Parent</td>
<td>9.65 (0.12)</td>
<td>2.91 (2.24)</td>
<td><a href="http://www.bristol.ac.uk/alspac/">www.bristol.ac.uk/alspac/</a></td>
<td>22,23</td>
</tr>
<tr>
<td>Generation R</td>
<td>2,211</td>
<td>CBCL/1.5-5</td>
<td>Parent</td>
<td>6.01 (0.38)</td>
<td>1.38 (1.69)</td>
<td><a href="http://www.generation.nl">www.generation.nl</a></td>
<td>24</td>
</tr>
<tr>
<td>GINI/USA</td>
<td>1,389</td>
<td>SDQ</td>
<td>Parent</td>
<td>10.04 (0.20)</td>
<td>2.71 (2.36)</td>
<td><a href="http://www.helmholtz-muenchen.de/epi/arbeitsgruppen/umweltepidemiologie/projects-projekte/lisa-plus/index.html">www.helmholtz-muenchen.de/epi/arbeitsgruppen/umweltepidemiologie/projects-projekte/lisa-plus/index.html</a></td>
<td>25,26</td>
</tr>
<tr>
<td>INMA</td>
<td>804</td>
<td>DSM based scale</td>
<td>Teacher</td>
<td>4.91 (0.69)</td>
<td>5.38 (6.83)</td>
<td><a href="http://www.proyectoinma.org/">www.proyectoinma.org/</a></td>
<td>27</td>
</tr>
<tr>
<td>MoBa</td>
<td>665</td>
<td>CBCL/1.5-5</td>
<td>Parent</td>
<td>3.05 (0.10)</td>
<td>2.05 (1.67)</td>
<td><a href="http://www.fhi.no/morogbarn">www.fhi.no/morogbarn</a></td>
<td>28</td>
</tr>
<tr>
<td>NTR</td>
<td>1,605</td>
<td>CBCL/6-12</td>
<td>Parent</td>
<td>9.95 (0.85)</td>
<td>3.25 (3.39)</td>
<td><a href="http://www.tweelingenregister.org">www.tweelingenregister.org</a></td>
<td>29</td>
</tr>
<tr>
<td>Raine</td>
<td>1,344</td>
<td>CBCL/6-12</td>
<td>Parent</td>
<td>10.58 (0.20)</td>
<td>2.60 (3.17)</td>
<td><a href="http://www.rainestudy.org.au">www.rainestudy.org.au</a></td>
<td>30-32</td>
</tr>
<tr>
<td>TEDS</td>
<td>2,606</td>
<td>Conners</td>
<td>Parent</td>
<td>7.88 (0.52)</td>
<td>10.51 (8.62)</td>
<td><a href="http://www.teds.ac.uk">www.teds.ac.uk</a></td>
<td>33</td>
</tr>
<tr>
<td>TRAILS</td>
<td>1,285</td>
<td>CBCL/6-12</td>
<td>Parent</td>
<td>11.08 (0.56)</td>
<td>4.27 (3.40)</td>
<td><a href="http://www.trails.nl">www.trails.nl</a></td>
<td>34</td>
</tr>
</tbody>
</table>

Note: ALSPAC = Avon Longitudinal Study of Parents and Children; CBCL = Child Behavior Checklist; GINI = German Infant Nutritional Intervention; INMA = Infancia y Medio Ambiente; USA = Influence of lifestyle factors on Development of the Immune System and Allergies in East and West Germany plus Air Pollution and Genetics on Allergy Development; MoBa = Norwegian Mother and Child Cohort Study; NTR = Netherlands Twin Register; SDQ = Strengths and Difficulties Questionnaire; TEDS = Twins Early Development Study; TRAILS = TRacking Adolescents’ Individual Lives Survey.
Phenotype
Different instruments were used across cohorts (Table 1), including the Attention Problems scale of the Child Behavior Checklist (CBCL) and the Teacher Report Form (TRF), the Hyperactivity scale of the Strengths and Difficulties Questionnaire (SDQ), and the DSM-IV ADHD items as, for example, included in the Conners Rating Scale (see Table S1, available online, for the items included in each scale).35-38 For the meta-analysis, one phenotype was selected from each cohort. Based on the phenotype that was most available, school-age ratings were chosen over preschool-age ratings, parent ratings over teacher ratings, and the measurement instrument with the largest information density was preferred over the other instruments (Conners DSM-IV > CBCL > SDQ).

SNP-Based Heritability
The variance in ADHD symptom scores accounted for by the SNPs was estimated using Genomic-Relationship-Matrix Restricted Maximum Likelihood (GREML) as implemented in the Genomic Complex Trait Analysis (GCTA) software.39,40 GREML is a linear mixed model that includes a genetic relatedness matrix (GRM) that contains a measure of genetic similarity between all possible pairs of (unrelated) individuals in a study. Genetic similarity is based on resemblance in SNP variants; hence the variance explained by the genetic relatedness matrix is often called the SNP heritability. Typically, only unrelated individuals (genetic relatedness <0.025) are included in the construction of a GRM to prevent the estimate of the SNP heritability to be biased upward.

GRM-based analyses were performed for the hyperactivity scale of the SDQ as measured in the Avon Longitudinal Study of Parents and Children (ALSPAC) at preschool (N = 5,510) and at school-age (N = 5,303), and for the Attention Problems scale of the CBCL 1.5-5 (N = 2,958) and the TRF (N = 1,901) measured in the Netherlands Twin Register (NTR) and Generation R cohorts. These analyses were not performed in the other cohorts because of the smaller sample sizes.

In ALSPAC, the GRM was constructed based on observed genotypes. Sex, age at measurement, and two principal components were included as fixed effects in the model. In NTR, the CBCL 1.5–5 was assessed when the children were 3 years of age and in Generation R when they were 6 years of age. The NTR and Generation R samples were combined to estimate the GRM. Individual-level genotype data from the NTR and Generation R were imputed together based on the Genome of the Netherlands reference set.41,42 Sex, age at measurement, sample, and five principal components were included as fixed effects in the model.

Data Quality Control and SNP and Gene-Based Association Meta-Analyses
Cohorts performed quality control (QC) and imputed their SNP genotype data using the March 2012 release of the 1000 Genomes reference set that includes all ethnicities.43 Each cohort ran their own optimal pre-imputation genotype QC. An overview of the pre-imputation QC metrics and imputation methods applied in each cohort is provided in Table S2, available online. Briefly, filtering on sample and SNP call rate was similar between cohorts, with the exception of the Twins Early Development Study (TEDS), which had a lower threshold. The thresholds for Hardy–Weinberg equilibrium, heterozygosity filtering, and other QC steps varied somewhat more. This likely has slightly decreased the final imputation quality of each cohort because prior SNP filtering does not improve the imputation quality, as losing genotyped SNPs only makes the imputation worse.44 Imputation with the Mach software has been shown to perform slightly better,44 but differences tend to be small. All filtering decisions will eventually result in fewer SNPs in the meta-analysis, bringing the risk of a loss of potential signal but not leading to false-positive results.

A linear regression of the phenotype on sex, age at measurement, genotype dose, and principal components was performed in all cohorts. All cohorts analyzed data from unrelated individuals, except for the Netherlands Twin Register (NTR), which included both twins from a dizygotic twin pair and corrected standard errors in PLINK —the —family —option (http://pngu.mgh.harvard.edu/ purcell/plink/).45 Table S2, available online, lists the analytic tools applied by each cohort.

Results were checked and meta-analyzed by two independent analysts. QC included calculation of the inflation factor lambda (the ratio of the observed versus the expected median \( \hat{\chi}^2 \)), format checking, visual inspection of QQ plots, Manhattan plots, histograms of minor allele frequency (MAF) and INFO scores, consistency of reported allele frequency with the reference set, consistency of reported \( p \) value with reported \( \beta \) value and standard error (SE), and consistency of reported SE with reported sample size, standard deviation (SD), and MAF. All files were filtered using the software EasyQC46 (www.genepi-regensburg.de/easyqc) based on R^2 metric >0.7 for MACH-based imputations and INFO metric >0.8 for IMPUTE-based imputations. This filter was applied to all SNPs to ensure conservatism. In addition, SNPs were filtered for MAF >0.03, Hardy–Weinberg equilibrium \( p \) value <10^-4, consistency of reported alleles and allele frequency with the reference set (maximum difference 0.2 with 1000G phase 1 v3), and duplicates (both occurrences removed).

As different phenotyping instruments were used across cohorts, the meta-analysis was based on \( p \) values and performed in METAL software (http://www.sph.umich.edu/csg/abecasis/metal/; option SCHEME SAMPLESIZE) including an application of genomic control to the results of the individual cohorts.47 Meta-analysis results were filtered on a total sample size >10,000. A \( p \) value of <5 \times 10^-8 was considered genome-wide significant.

Gene-based analyses were performed in MAGMA (Multi-marker Analysis of GenoMic Annotation).48 In MAGMA, a gene test-statistic is calculated as the mean of the \( \chi^2 \) statistics for all of the SNPs between the transcription start and stop sites of a gene. The gene \( p \) value is then obtained by using a known approximation of the sampling distribution. MAGMA corrects for gene size, number of SNPs in a gene, and linkage disequilibrium (LD) between SNPs in a gene using the SNP correlation matrix. We used the European ancestry samples from the 1000 Genomes project as reference data to estimate LD. Association was tested for 17,155 genes. The \( p \) value for gene-wide significance after Bonferroni correction was 0.05/17,155 = 2.91 \times 10^{-6}.

Comparing EAGLE and PGC ADHD Case-Control Meta-Analyses Results
We investigated the overlap in the results for the SNP and gene-based analyses obtained in EAGLE with the results of the PGC ADHD meta-analysis (P. Holmans for the PGC: presentation 21st World Congress of Psychiatric Genetics, October 2013, Boston, MA; http://2013.ispg.net/wp-content/uploads/2013/10/Oral-Presentations-Abstract-Book.pdf). The PGC sample comprised 5,621 cases and 13,589 controls. The overlap in SNP effects was investigated with SNP effect Concordance Analysis (SECA)49 and with Linkage Disequilibrium Score (LDSC) Regression analysis.50,51 Both methods require only the summary statistics of the association (meta-)analyses.
SECA takes the overlapping SNPs of both datasets and selects from each set of results SNPs with p values of ≤ .01, .05, .1, .2, .3, .4, .5, .6, .7, .8, .9, and 1.0. This leads to a 12 x 12 matrix indicating the overlap in SNP effects for each combination of SNP sets in the two analyses. SECA performs several tests based on these 144 cells. We report the results of the Fisher exact tests analyzing for each combination of the SNP sets whether the number of SNPs that are concordant in the direction of effects for the two phenotypes is above chance. Next, an empirical p value is calculated that indicates whether the overlap is higher than expected by chance given the multiple testing (144 tests).

After testing for SNP concordance by SECA, we also tested for overlap at the gene level, following the procedure described in Zhao et al.,52 in which gene-based associations were tested in GATES.53 Independent genes were then identified by examining the LD between the most strongly associated SNPs within each gene in the Genetic type I Error Calculator (GEC).54 Exact binomial statistical tests then determined whether the number of genes with p values < .01, < .05, or < .1 that were observed in both sets of results was significantly higher than expected.

The LDSC regression analysis was performed in the LDSC package (https://github.com/bulik/lodsc).50,51 This analysis yields SNP heritability estimates of the traits and a genetic correlation between the traits.

RESULTS

SNP-Based Heritability

The estimates of the SNP-based heritability are shown in Table 2. The estimates for the maternal ratings were 5% (not significant) for preschool SDQ and 13% (not significant) and 14% (p < .05) for preschool CBCL and school-age SDQ, respectively. For teacher ratings, an SNP-based heritability of 34% (p < .05) was observed. These results indicated that SNPs tag variants associated with various ADHD symptom scores.

SNP- and Gene-Based Meta-Analyses

The numbers of SNPs from each cohort that were included in the meta-analysis after QC are displayed in Table S3, available online. The QQ plot in Figure 1 shows the distribution of SNP p values from the meta-analysis filtered on SNPs that were present in at least 10,000 individuals. The lambda statistic of the meta-analysis was 0.98. Individual cohorts had lambdas of ≤ 1.08 (Table S3, available online), implying absence of population stratification. The Manhattan plot in Figure 2 shows that none of the SNPs reached genome-wide significance. However, the QQ plot shows some departure from the expected line for the smallest p values, which may reflect the polygenic nature of the trait, that is, many variants of small effects influencing ADHD symptoms. The strongest association was with rs6159542 on chromosome 19 (p = 1.48 x 10^-7). A summary of the top signals that crossed the threshold of suggestive association at p < 1 x 10^-5 is included in Table 3. Eight of the top nine variants were located in genes. As shown in the locus zoom plots55 of the top SNPs (Figure S1, available online), the signals of rs79846815, rs61227778, and rs77216358 were restricted to one to three SNPs, suggesting that they might not be genuine signals.

Table 4 shows the top 10 genes observed in the gene-based analysis as performed in MAGMA. Three genes, WASL, LMOD2, and ASB15, yielded gene-wide associations. An SNP in LMOD2 also yielded a suggestive association with ADHD symptom scores (Table 3). The locus zoom plot (Figure S2, available online) shows that several SNPs show similar signals in the three genes, as they are in a region with SNPs in high LD. Thus, the signals of the three genes are not independent of each other. The previously identified ADHD candidate genes did not show evidence of association. The p values ranged from 0.11 to 0.91 (Table S4, available online).

Overlap in Results Between the EAGLE and a PGC ADHD Case-Control Meta-Analysis

The SECA software tested for 144 combinations of SNP subsets obtained from the EAGLE and the PGC ADHD meta-analyses, whether there were more SNPs showing concordance in the direction of effects than expected by chance.

FIGURE 1 QQ-plot of all meta-analysis results based on at least 10,000 individuals.

Table 2

<table>
<thead>
<tr>
<th>Generation R/NTR</th>
<th>ALSPAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBCL (3 and 6 y)</td>
<td>TRF (7 y)</td>
</tr>
<tr>
<td>N</td>
<td>2,958</td>
</tr>
<tr>
<td>Variance explained</td>
<td>0.13</td>
</tr>
<tr>
<td>SE</td>
<td>0.11</td>
</tr>
<tr>
<td>p Value</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Note: CBCL = Child Behavior Checklist; SDQ = Strengths and Difficulties Questionnaire; TRF = Teacher Report Form.

Table 4 shows the top 10 genes observed in the gene-based analysis as performed in MAGMA. Three genes, WASL, LMOD2, and ASB15, yielded gene-wide associations. An SNP in LMOD2 also yielded a suggestive association with ADHD symptom scores (Table 3). The locus zoom plot (Figure S2, available online) shows that several SNPs show similar signals in the three genes, as they are in a region with SNPs in high LD. Thus, the signals of the three genes are not independent of each other. The previously identified ADHD candidate genes did not show evidence of association. The p values ranged from 0.11 to 0.91 (Table S4, available online).
chance (Fisher tests with an odds ratio of $\geq 1$ and a $p$ value of $\leq 0.05$). These analyses resulted in 111 SNP subsets with $p$ values of $<0.05$. This is significantly higher than expected by chance (empirical $p$ value = 0.001), clearly indicating an overlap between the SNPs associated with ADHD symptom scores and ADHD diagnoses. This was confirmed by the analysis investigating the overlap in genes rather than SNPs.52 The overlap in genes with a $p$ value of $<0.05$ or $>0.1$ in both cohorts was larger than expected by chance, with binomial $p$ values of 0.05 and $1.3 \times 10^{-3}$, respectively.

In the LDSC regression analysis, the genetic correlation between the EAGLE ADHD symptom scores and the PGC ADHD case-control phenotype was estimated at 0.96 (SE between the EAGLE ADHD symptom scores and the PGC both cohorts was larger than expected by chance, with evidence for a considerable common genetic background, with an estimate of the genetic correlation of 0.96).

The three associated genes, LMOD2 (7q31.32), ASB15 (7q31.32), and WASL (7q31.32), lie in a region with high LD, thereby making it difficult to decide on statistical grounds which gene contributes to this signal. Leiomodin (LMOD) is an actin-binding protein that acts as a filament nucleator in muscle cells.56 Ankyrin repeat and SOCS box containing 15 (ASB15) gene product belongs to the ASB family of proteins that are part of a ubiquination-mediated pathway.57 The ubiquitin proteasome pathway has also been suggested to play a role in adult ADHD.58 The protein encoded by Wiskott-Aldrich syndrome like (WASL) is involved in cytoskeletal organization during neuronal development, including synapse formation and neurite extension.59 Given the enrichment of genes involved in directed neurite outgrowth in the analysis of the combined results of published GWA studies on ADHD,60 WASL seems the most likely candidate to drive the signal.

A power analysis in Quanto61 suggested that the current sample has 80% power to detect a genome-wide significant effect explaining 0.21% of the variance, assuming an additive genetic effect, considering the largest possible sample of

![Manhattan plot of meta-analysis results based on at least 10,000 individuals.](image)

**FIGURE 2**

### DISCUSSION

The current study comprised the largest GWA analysis of continuous ADHD symptom scores in children to date. We found that common variants included in GWA studies explained variation in ADHD symptom scores assessed in the general population. The SNP heritability estimates for the various measures from the participating cohorts ranged from 0.05 to 0.34. The SNP heritability based on the results of the meta-analysis in all cohorts with a total of 17,666 children was estimated at 8%. We did not detect SNPs at genome-wide significance levels, but detected three genome-wide significant results in the gene-based analyses. The analyses investigating the overlap in genetic influences on ADHD symptoms scores and ADHD diagnosis provided evidence for a considerable common genetic background, with an estimate of the genetic correlation of 0.96.

### TABLE 3

Top Signals From Single Nucleotide Polymorphism (SNP) and Gene-Based Meta-Analyses: List of Independent Signals With $p < 1 \times 10^{-5}$

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Position (GRCh37)</th>
<th>Effect/Other Allele</th>
<th>Frequency (freq in refset)</th>
<th>Total N</th>
<th>Direction of Effect</th>
<th>z Score</th>
<th>p Value</th>
<th>Location in/to Nearest Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs56159542</td>
<td>19</td>
<td>19682971</td>
<td>T/C</td>
<td>0.21 (0.19)</td>
<td>17,666</td>
<td>$-$</td>
<td>-5.26</td>
<td>$1.48 \times 10^{-7}$</td>
<td>PBX4 intronic</td>
</tr>
<tr>
<td>rs46297772</td>
<td>7</td>
<td>152823816</td>
<td>A/G</td>
<td>0.93 (0.93)</td>
<td>16,322</td>
<td>$-$</td>
<td>-4.76</td>
<td>$1.97 \times 10^{-6}$</td>
<td>downstream ACTR38</td>
</tr>
<tr>
<td>rs79846815</td>
<td>7</td>
<td>134563570</td>
<td>A/T</td>
<td>0.97 (0.96)</td>
<td>11,175</td>
<td>$+$</td>
<td>4.75</td>
<td>$2.03 \times 10^{-6}$</td>
<td>CALD1 intronic</td>
</tr>
<tr>
<td>rs7809453</td>
<td>7</td>
<td>123301940</td>
<td>A/G</td>
<td>0.54 (0.56)</td>
<td>17,666</td>
<td>$-$</td>
<td>-4.69</td>
<td>$2.78 \times 10^{-6}$</td>
<td>LMOD2 exonic, synonymous</td>
</tr>
<tr>
<td>rs79162905</td>
<td>14</td>
<td>89796072</td>
<td>A/G</td>
<td>0.11 (0.10)</td>
<td>17,666</td>
<td>$-$</td>
<td>-4.68</td>
<td>$2.81 \times 10^{-6}$</td>
<td>FOXN3 intronic</td>
</tr>
<tr>
<td>rs146855089</td>
<td>2</td>
<td>77317636</td>
<td>A/G</td>
<td>0.26 (0.27)</td>
<td>17,666</td>
<td>$-$</td>
<td>-4.52</td>
<td>$6.18 \times 10^{-6}$</td>
<td>LRRTM4 intronic</td>
</tr>
<tr>
<td>rs10808119</td>
<td>7</td>
<td>101840716</td>
<td>A/G</td>
<td>0.46 (0.45)</td>
<td>17,666</td>
<td>$+$</td>
<td>4.50</td>
<td>$6.72 \times 10^{-6}$</td>
<td>CUX1 intronic</td>
</tr>
<tr>
<td>rs61227778</td>
<td>14</td>
<td>24578916</td>
<td>A/G</td>
<td>0.95 (0.96)</td>
<td>10,826</td>
<td>$+$</td>
<td>4.45</td>
<td>$8.62 \times 10^{-6}$</td>
<td>NRL intronic</td>
</tr>
<tr>
<td>rs77216358</td>
<td>11</td>
<td>120311157</td>
<td>A/G</td>
<td>0.96 (0.97)</td>
<td>11,042</td>
<td>$-$</td>
<td>-4.43</td>
<td>$9.48 \times 10^{-6}$</td>
<td>ARHGEF12 intronic</td>
</tr>
</tbody>
</table>

---

900 www.jaacap.org

JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY
VOLUME 55 NUMBER 10 OCTOBER 2016
N = 17,666, and assuming that a meta-analysis has as much statistical power as a single analysis of a similar sample size. The lack of genome-wide significant effects in combination with the observed SNP-based heritability estimates indicates that ADHD is likely to be highly polygenic, that is, influenced by many common genetic variants of small effect sizes. These results are in line with earlier studies on ADHD based on case-control samples. Following Yang et al., we estimated that the PGC ADHD subgroup GWA meta-analysis of 5,621 cases and 13,589 controls had 8% more power than the current study, and yet no significant results were found despite the SNP heritability estimated at 28% in this sample. Several other studies also found evidence for polygenicity of ADHD, with the exception of one study. The range of the SNP heritability estimates for the symptom scores in the current study (Table 2) seems quite large; however, these differences could well be due to chance, judging by the large standard errors.

The evidence provided for the genetic overlap in continuous and dichotomous measures of ADHD agrees with findings from previous studies with smaller samples. One of these studies also reported that an aggregate polygenic risk score derived from a sample of clinical cases of ADHD predicted preschool parent and school-age parent and teacher ratings of attention problems in a population-based cohort. This indicates common genetic variance across these measures and ages, which is confirmed by the SNP-based heritability of 8% that we calculated based on the results of the meta-analysis. Despite the use of various measures across the cohorts and the accompanying heterogeneity, there remains a signal after combining the results. Overall, it can be concluded that the different instruments assess an underlying common liability for ADHD. Therefore, combining various continuous ADHD measures assessed in the general population with dichotomous diagnosis of ADHD assessed in clinical samples can be a successful way to increase sample size and statistical power for GWA studies. This is supported by preliminary results of the PGC ADHD subgroup (R. Walters: presentation 23rd World Congress of Psychiatric Genetics, October 2015, Toronto, ON, Canada).

Efforts to decrease heterogeneity across studies by harmonizing phenotypes can result in a further increase in power to detect genetic effects. Behavioral genetic studies have reported that genetic factors are not entirely similar across instruments, raters, and ages. We ran additional SECA analyses to investigate the overlap in results of the current meta-analysis with the results from a GWAS in an independent sample of 727 Australian adolescents whose mothers provided retrospective ratings of their childhood attention skills and problems using the Strengths and Weaknesses of ADHD Symptom Scale (SWAN). These analyses did not show concordance in SNP effects, but we note that this could be due to the small size of the Australian sample. Statistical methods such as item response theory (IRT) can be used to synchronize different measurement instruments in a sophisticated manner, and have already been successfully applied in a GWA meta-analysis of personality measures.

Another way to refine the phenotype when longitudinal data are available is to test the effect of the SNP on a latent variable that reflects stability over time and is more heritable than a measure at a single time point.

To conclude, our results support the notion that ADHD is influenced by genes involved in neuronal development. By performing GWA meta-analyses in larger samples, we should be able to identify genetic variants for ADHD, further elucidating its biological foundation. The use of continuous ADHD symptom scores in population-based cohorts is an exciting possibility for achieving this goal.

TABLE 4 Top Signals From Single Nucleotide Polymorphism (SNP) and Gene-Based Meta-Analyses: Top 10 Genes From Gene-Based Tests in Multi-Marker Analysis of GenoMic Annotation (MAGMA)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>Start Position (GRCh37)</th>
<th>Stop Position (GRCh37)</th>
<th>No. of SNPs</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMOD2</td>
<td>7</td>
<td>123295861</td>
<td>123304147</td>
<td>13</td>
<td>1.46 x 10^-6</td>
</tr>
<tr>
<td>WASL</td>
<td>7</td>
<td>123321997</td>
<td>123389116</td>
<td>117</td>
<td>1.50 x 10^-6</td>
</tr>
<tr>
<td>ASB15</td>
<td>6</td>
<td>123249112</td>
<td>123277932</td>
<td>45</td>
<td>2.66 x 10^-6</td>
</tr>
<tr>
<td>CUX1</td>
<td>7</td>
<td>101459184</td>
<td>101927250</td>
<td>1003</td>
<td>6.03 x 10^-5</td>
</tr>
<tr>
<td>HAPLN4</td>
<td>19</td>
<td>19366450</td>
<td>19373596</td>
<td>13</td>
<td>6.10 x 10^-5</td>
</tr>
<tr>
<td>CILP2</td>
<td>19</td>
<td>19649074</td>
<td>19657468</td>
<td>16</td>
<td>6.35 x 10^-5</td>
</tr>
<tr>
<td>LRRM4</td>
<td>2</td>
<td>76974849</td>
<td>77749502</td>
<td>1858</td>
<td>8.50 x 10^-5</td>
</tr>
<tr>
<td>ZNF234</td>
<td>19</td>
<td>44645710</td>
<td>44666462</td>
<td>19</td>
<td>8.73 x 10^-5</td>
</tr>
<tr>
<td>NDUFA13</td>
<td>19</td>
<td>19627019</td>
<td>19639013</td>
<td>19</td>
<td>1.09 x 10^-4</td>
</tr>
<tr>
<td>RWDD4</td>
<td>4</td>
<td>184560789</td>
<td>184580331</td>
<td>41</td>
<td>1.52 x 10^-4</td>
</tr>
</tbody>
</table>

Note: Boldface data denotes gene-wide significance (p < 2.91 x 10^-8). Chr = chromosome.
Dr. St. Pourcain is with MRC Integrative Epidemiology Unit (MRC IEU), University of Bristol, UK, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, and School of Experimental Psychology, University of Bristol. Dr. Greven is with Donors Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Karakter, Child and Adolescent Psychiatry University Center, Radboud University Nijmegen, and MRC Social Genetic and Developmental Psychiatry Centre, King’s College London. Dr. Pappas is with Generation R Study Group, and Pedagogical and Education Sciences, Erasmus University Rotterdam, The Netherlands. Drs. Tiessler and Thiering are with Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neufahrn, Germany and the Division of Metabolic and Nutritional Medicine, Munich, and Dr. van Haaren Children’s Hospital, University of Munich Medical Center, Germany. Mr. Avg. Ms. Wang, and Dr. Pennell are with School of Women’s and Infants’ Health, University of Western Australia, Perth. Dr. Nolte is with University of Groningen, University Medical Center Groningen, The Netherlands. Ms. Vilar-Tejedor is with Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Universitat Pompeu Fabra (UPF), Barcelona, and CIBER Epidemiology and Public Health (CIBERESP), Madrid. Mr. Bolecis is with Gothenburg University, Sweden. Drs. Ebeger, Martis, and Medland are with QIMR Berghofer Medical Research Institute, Brisbane, Australia. Drs. Zhao and Nyholt are with Institute of Health and Biomedical Innovation, Queensland University of Technology, Queensland, Australia. Drs. Davies and Ehr are with Avera Institute for Human Genetics, SD. Drs. Evans, Kemp, and Ring are with MRC IEU, School of Social and Community Medicine, and School of Social and Community Medicine, University of Bristol, and Diamantina Institute, Translational Research Institute, University of Queensland, Brisbane. Ms. Fedlova is with Biological Psychology, VU University Amsterdam. Dr. Guerini is with CREAL, UPF, CIBERESP, and Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Center, Sophia Children’s Hospital, The Netherlands. Dr. Hattenga is with Biological Psychology, VU University, and EMGO+- Institute for Health and Care Research, VU University Medical Center. Dr. Hulsbergen is with Vermont Center for Children, Youth and Families, and College of Medicine, University of Vermont, Burlington, and Child and Adolescent Psychiatry, Erasmus Medical Center. Drs. Jigues, Myhr, and Stoltenberg are with the Norwegian Institute of Public Health, Oslo. Ms. Kapatil and Drs. Tzitzikos and Pliam are with MRC Social, Genetic and Developmental Psychiatry Centre, King’s College London. Mr. Murcia is with CIBERESP, and FISABIO—Universitat Jaume I—Universitat de València Joint Research Unit of Epidemiology and Environmental Health, Valencia, Spain. Drs. Ornel and Hartman are with the Interdisciplinary Center Psychopathology and Emotion regulation (ICPE), University Medical Center Groningen. Drs. Standl and Heinrich are with Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neufahrn, Germany. Drs. Stegiakouli and Timpector are with MRC IEU, Dr. Timpector is also with School of Social and Community Medicine, University of Bristol. Dr. van der Most is with University of Groningen and University Medical Center Groningen. Dr. Vanhoutte is with Program in Medical and Population Genetics and Stanley Center for Psychiatric Genetics, Broad Institute of Massachusetts Institute of Technology, Boston, Analytic and Translation Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, and Harvard University, Cambridge, MA. Dr. Jacobsson is with Obstetrics and Gynecology, Gothenburg University, and the Norwegian Institute of Public Health. Dr. Sunyer is with CREAL, IRM (Hospital del Mar Medical Research Institute), Barcelona, UPF, and CIBERESP. Dr. Whitehouse is with Telethon Kids Institute, University of Western Australia, Perth. Dr. Davey Smith is with MRC IEU, and School of Social and Community Medicine. Dr. Tiemeier is with Epidemiology, Child and Adolescent Psychiatry, and Psychiatry, Erasmus Medical Center. Dr. Posthuma is with the Generation R Study Group, Erasmus MC Rotterdam, the Netherlands, Child and Adolescent Psychiatry, Erasmus Medical Center, Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University, and Clinical Genetics, University Medical Center, Dr. Boomsma is with Biological Psychology, VU University, Neuroscience Campus Amsterdam, VU University, and EMGO+- Institute for Health and Care Research, VU University Medical Center.

ALSPAC: The UK Medical Research Council and the Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and they will serve as guarantors for the contents of this paper. GWAS data were generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (laboratory Corporation of America) using support from 23andMe.

Generation R: The Generation R Study is made possible by financial support from the Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam, the Netherlands Organisation for Health Research and Development (ZonMw), the Netherlands Organisation for Scientific Research (NWO), and the Ministry of Health, Welfare and Sport. H.T. received additional grants from the Netherlands Organisation for Health Research and Development (ZonMw VIDI 017.106.370). The work of A.H. and D.P. was supported by a grant from the Dutch Science Organisation for Scientific Research (NWO) 433-09-228 and 453-14005.

GINI/USA: Personal and financial support by the Munich Center of Health Sciences (MCHEALTH) as part of the Ludwig-Maximilians University Munich LMU innovative is gratefully acknowledged.

INWA: This study was funded by grants from the Spanish Instituto de Salud Carlos III (IB06/02/0041, G03/176, FIS PI041436, PI081151, PI041705, PI061756, PI091958, and PS09/00432, FIS-FEDER 03/1615, 04/1509, 04/1112, 04/1931, 05/1079, 05/1052, 06/1213, 07/0314, 09/02647 11/01007, 11/02591, 11/02038, 13/1944, 13/2032, CP11/0178 and MS13/00054), Spanish Ministry of Science and Innovation (SAF200800357), European Commission (ENGAGE project and grant agreement HEALTH-F4-2007-201413, HEALTH-2010-2.4.5-1, FP7-ENV-2011-1c 282957), Fundación La Marató de TV3, Generalitat de Catalunya-CIBER 199595G 00241, and Conselleria de Sanitat Generalitat Valenciana. Part of the DNA extractions and genotyping was performed at the Spanish Genotyping Centre (CEGENBarcelona). N. Vilar-Tejedor thanks the Agència de Gestió d’Ajuts Universitaris i de Recerca – Generalitat de Catalunya for her predoctoral grant (2015 FI_B_00636).

MoBa (Mother and Child Cohort of NPH): This work was supported by grants from the Norwegian Research Council (FUGE 183220/S10, FRIMEDBIO 05ES23601), Swedish Medical Society (SLS 200821198), Jane and Dan Olsson Foundations, and Swedish government grants to researchers in the public health service (ALFGBG-2863, ALFGBG-11522, ALFGBG-426411), and the European Community’s Seventh Framework Programme (FP7/2007–2013), ENGAGE Consortium, grant agreement HEALTH-F4-2007-201413. The Norwegian Mother and Child Cohort Study was also supported by the Norwegian Ministry of Health and the Ministry of Education and Research, NIH/NIAMS (contract no. 1RC2-MD006971), NSF (grant no. 0423280), NIH/NIAMS (grant no. 1U01 NS 047537-01 and grant no. 2 U01 NS 047537-06A1), and the Norwegian Research Council/FUGE (grant no. 151918/S10). Researchers interested in using MoBa data must obtain approval from the Scientific Management Committee of MoBa and from the Regional Committee for Medical and Health Research Ethics for access to data and biological material. Researchers are required to follow the terms of an Assistance Agreement containing a number of clauses designed to ensure protection of privacy and compliance with relevant laws. For further information, contact the principal investigator of MoBa, Per Magnus (per.magnus@rhi.no).

The Netherlands Twin Register: Genomics of Mental Illness (European Research Council ERC-230374, Genetic influences on stability and change in psychopathology from childhood to young adulthood (ZonMW 91210020), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI–NL 184.021.007); VU University’s Institute for Health and Care Research (EMGO+-) and Neuroscience Campus Amsterdam (NCA; Community’s Seventh Framework Program (FP7/2007-2013); ENGAGE [HEALTH-F4-2007-201413], the Avera Institute, Sioux Falls, South Dakota, USA, and Grand Opportunity (grants 1RC2 MH098951 and 1RC2 MH098955). Raine: The authors gratefully acknowledge the NHMRC for their long-term contribution to funding the study over the last 25 years and also the following Institutions for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA), Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health Sciences, Telethon Kids Institute and Women and Infants Research Foundation (King Edward Memorial Hospital), Curtin University, and Edith Cowan University. The authors gratefully acknowledge the support of the Western Australian Genetic Epidemiology Resource and the Western Australian DNA Bank (both National Health and Medical Research Council of Australia National Enabling Facilities). The authors also acknowledge the support of the Healthway-Healthway West Australia, the National Health and Medical Research Council of Australia (Grant 572613), and the Canadian Institutes of Health Research (Grant MOP 82893). The authors gratefully acknowledge the assistance of the Wind Over Water Foundation, the Telethon Institute for Child Health Research, and the Raine Medical Research Foundation of the University of Western Australia. A.J.O’W. was supported by a Senior Research Fellowship from the NHMRC.
GWAS META-ANALYSIS OF PEDIATRIC ADHD SYMPTOMS

[Grant number 1077966]. This work was supported by resources provided by the Pavey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

TEDS: The Twins Early Development Study (TEDS) is supported by a program grant to R.P. from the UK Medical Research Council [G0501245], and previously G0500079, with additional support from the US National Institutes of Health [HD044454, HD059215]. R.P. is supported by a Medical Research Council Research Professorship award [G19/2] and a European Research Council Advanced Investigator award [295366]; M.T. is supported by British Academy Postdoctoral Fellowship [G10001F]; E.K. is supported by the Institute of Psychiatry Excellence/Medical Research Council postgraduate Studentship.

This research is part of the Tracking Adolescents’ Individual Lives Survey (TRAILS). Participating centers of TRAILS include the University Medical Center and University of Groningen, the Erasmus University Medical Center Rotterdam, the University of Utrecht, the Radboud Medical Center Nijmegen, and the Parnassia Bavo group, all in the Netherlands. TRAILS has been financially supported by various grants from the Netherlands Organization for Scientific Research NWO (Medical Research Council program grant GB-370 940-38011; ZonMW Brainpower grant 1000-0104, ZonMW Risk Behavior and Dependence grants 060-060007-1118, ZonMW Culture and Health grant 26198710, Social Sciences Council medium-sized investment grants GB-MaGW 48001-006 and GB-MaGW 48007001; Social Sciences Council project grants GB-MaGW 45204-314 and GB-MaGW 45206-004), NWO large-sized investment grant 175.010.2003.005; NWO longitudinal Survey and Panel funding 48108013 and 48111001), the Dutch Ministry of Justice (WODC), the European Science Foundation (EuroSTRESS project FP006), Biobanking and Biomolecular Resources Research Infrastructure BBMRI-NL (CP 32), and the participating universities.

EAGLE Working Groups and EArly Genetics and Lifecourse Epidemiology (EAGLE) Consortium (http://www.wikigenes.org/e/art/e/348.html): EAGLE Working Groups and EArly Genetics and Lifecourse Epidemiology (EAGLE) Consortium (http://www.wikigenes.org/e/art/e/348.html). The authors would particularly like to thank all the participants for their generous collaboration. A full roster of the INMA Project Investigators can be found at http://www.proyectoima.org/presentacion-inma/listado-investigadores.html. MoBa (Mother and Child Cohort of NPH): The authors are grateful to everyone who participated in this research or worked on this project to make it possible.

Disclosure: Dr. Hudziak has received grant or research support from the National Institutes of Health, the National Institute of Mental Health, the National Institute of Diabetes and Digestive and Kidney Diseases, and the National Institute of Environmental Health. His primary appointment is with the University of Vermont. He has additional appointments with Erasmus University Rotterdam, Netherlands, Washington University School of Medicine in St. Louis, Missouri, and the Giesel School of Medicine at Dartmouth in Hanover, New Hampshire. Dr. Middeldorp, GroenBKrhuis, St Piorcarn, Groen, Poppa, Tiemeier, Nolke, Ebejer, Zhao, Davies, Ehli, Evans, Ouezens, Hotten, Jukes, Martin, Myhre, Cemel, Ring, Standl, Stierakiouli, Stoltenberg,Tierping, Trzaskowski, van der Most, Nyholt, Medland, Neale, Jacobson, Sunyar, Hartman, Whitehouse, Plomin, Smith, Tiemeier, Posthuma, Buitelaar, and M. Hammerschlag report no biomedical financial interests or potential conflicts of interests.

Correspondence to Christel M. Middeldorp, MD, PhD, VU University Amsterdam, Biological Psychology, Van der Boechorststraat 1, 1081 BT, Amsterdam, Netherlands. e-mail: c.m.middeldorp@vum.nl

08908567/$36.00 ©2016 American Academy of Child and Adolescent Psychiatry

http://dx.doi.org/10.1016/j.jaac.2016.05.025

JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY VOLUME 55 NUMBER 10 OCTOBER 2016

903 www.jaacap.org
REFERENCES


32. Williams LA, Evans SF, Newnham JP. Prospective study of factors influencing the relative weights of the placenta and the newborn infant. BMJ. 1997;314:1864-1868.


54. Li MX, Yeung JM, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. Hum Genet. 2012;131:747-756.


FIGURE S1  Locus zoom plots of the nine single nucleotide polymorphisms (SNPs) with suggestive association at $p < 1 \times 10^{-5}$.  

FIGURE S1  Continued
FIGURE S1 (continued).

FIGURE S2 Locus zoom plot of the region with the three genewide significant signals. Note: SNPs = single nucleotide polymorphisms.
TABLE S1  Item Content of Attention-Deficit/Hyperactivity Disorder (ADHD) Symptom Scales Included in the Genome-Wide Association (GWA) Meta-Analysis

<table>
<thead>
<tr>
<th>Scale</th>
<th>Item Content</th>
</tr>
</thead>
</table>
| CBCL 1.5—5: Attention Problems scale | - Can’t concentrate, can’t pay attention for long  
- Can’t sit still, restless, or hyperactive  
- Poorly coordinated or clumsy  
- Quickly shifts from one activity to another  
- Wanders away |
| CBCL 6—18: Attention Problems scale | - Acts too young for his/her age  
- Fails to finish things he/she starts  
- Can’t concentrate, can’t pay attention for long  
- Can’t sit still, restless, or hyperactive; confused or seems to be in a fog  
- Daydreams or gets lost in his/her thoughts  
- Impulsive or acts without thinking  
- Poor school work  
- Inattentive or easily distracted  
- Stares blankly |
| SDQ: Hyperactivity-Inattention scale | - Restless, overactive, cannot stay still for long  
- Constantly fidgeting or squirming  
- Easily distracted, concentration wanders  
- Thinks things out before acting  
- Sees tasks through to the end, good attention span |
| Conners Rating Scales—Revised: Long Form; Other DSM-IV—based rating scales* | - Often fails to give close attention to details or makes careless mistakes in schoolwork, work, or other activities  
- Often has difficulty sustaining attention in tasks or play activities  
- Often does not seem to listen when spoken to directly  
- Often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (not due to oppositional behavior or failure of comprehension)  
- Often has difficulty organizing tasks and activities  
- Often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (such as schoolwork or homework)  
- Often loses things necessary for tasks or activities at school or at home (e.g., toys, school assignments, pencils, books, or tools)  
- Is often easily distracted by extraneous stimuli  
- Is often forgetful in daily activities  
- Often fidgets with hands or feet or squirms in seat  
- Often leaves seat in classroom or in other situations in which remaining seated is expected  
- Often runs about or climbs excessively in situations in which it is inappropriate  
- Often has difficulty playing or engaging in leisure activities quietly  
- Is often “on the go” or often acts as if “driven by a motor”  
- Often talks excessively  
- Often has difficulty awaiting turn  
- Often blurts out answers to questions before they have been completed  
- Often interrupts or intrudes on others, e.g., butts into other children’s games |

Note: CBCL = Child Behavior Checklist; SDQ = Strengths and Difficulties Questionnaire.

*Items may be phrased slightly differently across scales.
<table>
<thead>
<tr>
<th>Cohort</th>
<th>Genotyping Platform</th>
<th>Call rate</th>
<th>MAF</th>
<th>HWE</th>
<th>Other filters</th>
<th>Call rate</th>
<th>Heterozygosity</th>
<th>Ethnicity</th>
<th>Sex mismatches</th>
<th>Other filters</th>
<th>Imputation Software</th>
<th>Post-Imputation Filters</th>
<th>Association Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALSPAC</td>
<td>Illumina HumanHap550 quad-chip</td>
<td>0.95</td>
<td>0.01</td>
<td>5E-7</td>
<td></td>
<td>0.97</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>&gt;10% Identity by descent, insufficient sample replication</td>
<td>Minimac and Mach</td>
<td>None</td>
<td>Mach2Q TLV112</td>
</tr>
<tr>
<td>Generation R</td>
<td>Illumina Human 610 and 660 Quad Array Affymetrix 5.0 and Affymetrix 6.0</td>
<td>0.95</td>
<td>0.001</td>
<td>1E-7</td>
<td></td>
<td>0.975</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Relatedness</td>
<td>Minimac and Mach</td>
<td>None</td>
<td>Plink 1.07</td>
</tr>
<tr>
<td>GINI/ LISA</td>
<td>Affymetrix 5.0 and Affymetrix 6.0</td>
<td>0.95</td>
<td>0.01</td>
<td>1E-5</td>
<td></td>
<td>0.95</td>
<td>&gt; 4 SD</td>
<td>Yes</td>
<td>Yes</td>
<td>SNIPTEST NA for BETA, SE and PVAL</td>
<td>SNPTEST v2.4.1</td>
<td>SNPTEST v2.4.1</td>
<td></td>
</tr>
<tr>
<td>INMA</td>
<td>Illumina Human Omni I</td>
<td>0.95</td>
<td>0.01</td>
<td>1E-6</td>
<td></td>
<td>0.98</td>
<td>No</td>
<td>Yes</td>
<td>LRR SD &gt; 0.3, duplicates, relatedness</td>
<td>SHAPEIT (v2_r6644), Impute (version 2.3.0)</td>
<td>None</td>
<td>SNPTest v2.5-beta4</td>
<td></td>
</tr>
<tr>
<td>MoBa</td>
<td>Illumina Human 660W Quad Array</td>
<td>0.97</td>
<td>0.01</td>
<td>1E-6</td>
<td>Mitochondrial SNPs, chrY + PAR SNPs, SNPs that could not be updated to hg37, non-“rs” SNPs</td>
<td>0.96</td>
<td>&gt; 4 SD</td>
<td>Yes</td>
<td>Yes</td>
<td>Relatedness</td>
<td>Minimac and Mach</td>
<td>None</td>
<td>Plink 1.07</td>
</tr>
<tr>
<td>NTR</td>
<td>Affymetrix 6.0</td>
<td>0.95</td>
<td>0.01</td>
<td>1E-5</td>
<td>Double-typed error rate &gt; 0.02, Mendel error rate &gt; 0.02, allele frequency difference with reference set &gt; 0.20, C/G and A/T SNPs with MAF &gt; 0.35</td>
<td>0.90</td>
<td>F &gt; 0.10 or F &lt; -0.10</td>
<td>Yes</td>
<td>IBS/IBD discrepancies, Mendel error rate &gt; 0.02</td>
<td>Minimac and Mach</td>
<td>Plink NA for BETA, SE and PVAL</td>
<td>Plink 1.07</td>
<td></td>
</tr>
<tr>
<td>Raine</td>
<td>Illumina Human 660W Quad Array</td>
<td>0.95</td>
<td>0.01</td>
<td>5.7E-7</td>
<td>C/G and A/T SNPs removed</td>
<td>0.97</td>
<td>F &gt; 0.1875; heterozygosity &gt; 0.30</td>
<td>Yes</td>
<td>Yes</td>
<td>Mach</td>
<td>Minimac and Mach</td>
<td>None</td>
<td>Plink 1.07</td>
</tr>
<tr>
<td>Cohort</td>
<td>Genotyping Platform</td>
<td>Pre-Imputation Variant Filters</td>
<td>Pre-Imputation Sample Filters</td>
<td>Post-Imputation Filters</td>
<td>Association Software</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------------</td>
<td>-------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------</td>
<td>----------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEDS</td>
<td>Affymetrix 6.0</td>
<td>Call rate 0.80, MAF 0.01, HWE 1E-20, Other filters SNPTEST info &gt; 0.975</td>
<td>Call rate 0.98, Heterozygosity yes, Ethnicity yes, Sex mismatches yes, Other filters Relatedness (IBD &lt; 5%), regenotyping low concordance, Duplicates</td>
<td>Impute v2</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRAILS</td>
<td>Illumina CytoSNP12 v2</td>
<td>Call rate 0.95, MAF 0.01, HWE 1E-3, Other filters chr X SNPs &gt; 1% heterozygous, Other filters Callrate</td>
<td>Callrate</td>
<td>Impute v2</td>
<td>SNPtest</td>
<td>10%, duplicates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: ALSPAC = Avon Longitudinal Study of Parents and Children; GINI = German Infant Nutritional Intervention; HWE = Hardy Weinberg equilibrium; IBD = identity by descent; IBS = identity by state; INMA = INfancia y Medio Ambiente; ISA = Influence of Life-style factors on Development of the Immune System and Allergies in East and West Germany plus Air Pollution and Genetics on Allergy Development; IRR = log r ratio; MAF = minor allele frequency; MDS = multidimensional scaling; MoBa = Norwegian Mother and Child Cohort Study; NTR = Netherlands Twin Register; QC = quality control; SNP = single nucleotide polymorphism; TEDS = Twins Early Development Study; TRAILS = TRacking Adolescents’ Individual Lives Survey.
### TABLE S3  Results of the Data Cleaning for the Nine Cohorts Included in the Meta-Analysis

<table>
<thead>
<tr>
<th>Cohort</th>
<th>N</th>
<th>No. of Variants Uploaded</th>
<th>No. of Variants Cleaned</th>
<th>Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALSPAC</td>
<td>5,757</td>
<td>31,326,386</td>
<td>5,942,106</td>
<td>1.01</td>
</tr>
<tr>
<td>Generation R</td>
<td>2,211</td>
<td>30,072,738</td>
<td>5,907,888</td>
<td>1.02</td>
</tr>
<tr>
<td>GINI/USA</td>
<td>1,389</td>
<td>16,275,553</td>
<td>5,554,016</td>
<td>1.02</td>
</tr>
<tr>
<td>INMA</td>
<td>804</td>
<td>16,105,103</td>
<td>6,245,251</td>
<td>1.08</td>
</tr>
<tr>
<td>MOBA</td>
<td>665</td>
<td>14,154,076</td>
<td>6,177,049</td>
<td>1.02</td>
</tr>
<tr>
<td>NTR</td>
<td>1,605</td>
<td>8,868,990</td>
<td>5,654,673</td>
<td>1.03</td>
</tr>
<tr>
<td>Raine</td>
<td>1,338</td>
<td>28,625,631</td>
<td>5,260,671</td>
<td>0.99</td>
</tr>
<tr>
<td>TEDS</td>
<td>2,606</td>
<td>12,223,562</td>
<td>5,572,678</td>
<td>0.98</td>
</tr>
<tr>
<td>TRAILS</td>
<td>1,285</td>
<td>18,183,428</td>
<td>5,763,633</td>
<td>1.02</td>
</tr>
</tbody>
</table>

Note: ALSPAC = Avon Longitudinal Study of Parents and Children; GINI = German Infant Nutritional Intervention; INMA = Infancia y Medio Ambiente; ISA = Influence of Lifestyle factors on Development of the Immune System and Allergies in East and West Germany plus Air Pollution and Genetics on Allergy Development; MoBa = Norwegian Mother and Child Cohort Study; NTR = Netherlands Twin Register; TEDS = Twins Early Development Study; TRAILS = TRacking Adolescents’ Individual Lives Survey.

### TABLE S4  Results of Gene-Based Tests for Previously Identified Attention-Deficit/Hyperactivity Disorder (ADHD) Candidate Genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>Start Position (GRCh37)</th>
<th>Stop Position (GRCh37)</th>
<th>No. of SNPs</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRD4</td>
<td>11</td>
<td>637305</td>
<td>640706</td>
<td>5</td>
<td>.88</td>
</tr>
<tr>
<td>DRD5</td>
<td>4</td>
<td>9783258</td>
<td>9785633</td>
<td>2</td>
<td>.84</td>
</tr>
<tr>
<td>GIT1</td>
<td>17</td>
<td>27900487</td>
<td>27916610</td>
<td>16</td>
<td>.60</td>
</tr>
<tr>
<td>HTR1B</td>
<td>6</td>
<td>78171948</td>
<td>78173120</td>
<td>2</td>
<td>.25</td>
</tr>
<tr>
<td>NOS1</td>
<td>12</td>
<td>117465947</td>
<td>117799607</td>
<td>347</td>
<td>.18</td>
</tr>
<tr>
<td>SLC6A3</td>
<td>5</td>
<td>1392905</td>
<td>1445543</td>
<td>137</td>
<td>.50</td>
</tr>
<tr>
<td>SLC6A4</td>
<td>17</td>
<td>28523376</td>
<td>28562954</td>
<td>48</td>
<td>.91</td>
</tr>
<tr>
<td>SNAP25</td>
<td>20</td>
<td>10199477</td>
<td>10288065</td>
<td>173</td>
<td>.88</td>
</tr>
<tr>
<td>SLC9A9</td>
<td>3</td>
<td>142984064</td>
<td>143567373</td>
<td>1457</td>
<td>.11</td>
</tr>
</tbody>
</table>

Note: Chr = chromosome; SNP = single nucleotide polymorphism.