# Attention-Deficit/Hyperactivity Disorder Polygenic Risk Scores Predict Attention Problems in a Population-Based Sample of Children

Maria M. Groen-Blokhuis, MD, PhD, Christel M. Middeldorp, MD, PhD, Kees-Jan Kan, PhD, Abdel Abdellaoui, MSc, Catharina E.M. van Beijsterveldt, PhD, Erik A. Ehli, PhD, Gareth E. Davies, PhD, Paul A. Scheet, PhD, Xiangjun Xiao, MSc, James J. Hudziak, MD, Jouke-Jan Hottenga, PhD, Psychiatric Genomics Consortium ADHD Working Group, Ben M. Neale, PhD, Dorret I. Boomsma, PhD

**Objective:** Clinically, attention-deficit/hyperactivity disorder (ADHD) is characterized by hyperactivity, impulsivity, and inattention and is among the most common childhood disorders. These same traits that define ADHD are variable in the general population, and the clinical diagnosis may represent the extreme end of a continuous distribution of inattentive and hyperactive behaviors. This hypothesis can be tested by assessing the predictive value of polygenic risk scores derived from a discovery sample of ADHD patients in a target sample from the general population with continuous scores of inattention and hyperactivity. In addition, the genetic overlap between ADHD and continuous ADHD scores can be tested across rater and age. **Method:** The Psychiatric Genomics Consortium has performed the largest genome-wide analysis (GWA) study of ADHD so far, including 5,621 clinical patients and 13,589 controls. The effects sizes of single nucleotide polymorphisms (SNPs) estimated in this meta-analysis were used to obtain individual polygenic risk scores in an independent population-based cohort of 2,437 children from the Netherlands Twin Register. The variance explained in Attention Problems (AP) scale scores by the polygenic risk scores was estimated by linear mixed modeling. **Results:** The ADHD polygenic risk scores significantly predicted both parent and teacher ratings of AP in preschool- and school-aged children. **Conclusion:** These results indicate genetic overlap between a diagnosis of ADHD and AP scale scores across raters and age groups and provides evidence for a dimensional model of ADHD. Future GWA studies on ADHD can likely benefit from the inclusion of population-based cohorts and the analysis of continuous scores. J. Am. Acad. Child Adolesc. Psychiatry, 2014;53(10):1123–1129. **Key Words:** ADHD, attention problems, polygenic scores, genetics, dimensional models

Attention-deficit/hyperactivity disorder (ADHD) is a condition characterized by age-inappropriate hyperactivity/impulsivity and inattention, resulting in significant impairment in about 5% of children.[1,2] In the diagnostic manuals used in clinical practice, for example, the *International Classification of Diseases, 10th Revision (ICD-10)*, *DSM-IV*, and the *DSM-5*,[1,3,4] a clinical diagnosis of ADHD is a binary trait that can be useful for guiding treatment and care. At the population level, ADHD may represent the extreme end of a continuous distribution of inattentive and hyperactive behaviors.[5-7] Classical twin studies support the validity of the dimensional model,[8] but at this point in time, an additional approach to test for a dimensional model of ADHD is to assess whether genetic risk factors for an ADHD diagnosis influence behavior across the entire spectrum of inattentive and hyperactive behavior. Such an approach also may clarify apparent differences in the ADHD

Supplemental material cited in this article is available online.

assessment across raters and age groups. Correlations between parent and teacher ratings are generally only moderate, as are correlations within maternal ratings across preschool and school age.[8-10] Previous studies indicate that the extent to which assessment from different raters and across different ages overlap is due to overlap of genetic effects across both raters and time,[8-12] but these results are based on, for example, latent variable modeling approaches rather than on measured genetic variants. ADHD diagnoses and continuous measures of ADHD behaviors are highly heritable in childhood, with about 60% to 80% of the variance due to genetic factors.[13-17] Despite this high heritability, current genome-wide association (GWA) studies have thus far been unsuccessful in detecting genetic risk variants for ADHD at genome-wide significant levels, suggesting a high degree of polygenic inheritance.[18] A study by the Psychiatric Genomics Consortium (PGC) showed that 28% of the liability to ADHD is explained by single nucleotide polymorphisms (SNPs) present on platforms that are commonly used for genome-wide genotyping.[19] These observations imply that many common variants of small effect stay undetected in current GWA studies due to limited sample size but very likely contribute to the genetic liability of ADHD. The effect sizes obtained in ADHD GWA studies can be used to estimate the genetic risk of the individual; so-called polygenic risk scores are obtained by multiplying the measured number of risk alleles at a particular locus by the effect size observed in a GWA study summing over all SNPs that surpass a certain threshold of significance.[20,21] With regard to ADHD, polygenic risk scores based on the results of the PGC ADHD meta-analysis published in 2010 significantly predicted ADHD status in an independent sample of 452 clinical patients with ADHD and 5,081 controls, with higher polygenic risk scores in patients with ADHD and comorbid aggression.[18,22] Polygenic risk scores can also be used to assess the genetic overlap across traits. For example, polygenic risk scores based on a GWA study on schizophrenia predict quantitative measures of psychosis.[23] Similarly, polygenic risk scores based on a GWA study in patients with major depressive disorder (MDD) are predictive of continuous scores of anxiety and depression in a general population sample.[24] In the current study, we obtained polygenic risk scores to assess the genetic overlap between clinically assessed ADHD and attention problems (AP) in a general population sample of children who were rated by their parent at preschool age and by their parents and teachers at school age.

## METHOD

Genotype and phenotype data were available in a sample of 2,437 children of Dutch descent who are registered with the Netherlands Twin Register (NTR).[25,26] In the Young NTR (YNTR), surveys assessing the health and behavior of newborn twins are sent out to their parents at registration and at age 2, 3, 5, 7, 10, and 12 years. At age 7, 10, and 12 years, parents are asked for their consent to invite the teachers of the twins to provide ratings of the children's behavior.

### AP

Age-appropriate versions of the Achenbach System of Empirically Based Assessment (ASEBA) have been included in the YNTR surveys.[27,28] At ages 3, 7, 10, and 12 years the Child Behavior Checklist (CBCL) was collected from parents. At ages 7, 10, and 12 years, the Teacher Report Form (TRF) was included in teacher surveys. Respondents were asked to rate the child's behavior on ~120 items on a 3-point scale (0 = not true; 1 = somewhat or sometimes true; 2 = very true or often true). The AP scale describes hyperactive and inattentive behavior. The AP scale contains 5 items at preschool age, and at school age, 10 items for parents, and 26 items for teachers. When multiple measures were available for the school-age (age 6–13 years) mother or teacher ratings, the measure closest to age 10 was chosen. There were 2,132 twins with maternal AP ratings at school age; for 1,888 twins (89%), AP was assessed between age 9 and 11, for 50 twins at age 12, and for 194 twins, at age 7 or 8 years. Teacher ratings were available at age 9 to 11 for 1,018 twins, at age 7 to 8 for 152 twins, and at age 12 for 442 twins. Maternal and paternal ratings were highly correlated (r = 0.71 and 0.73 for preschool and school age) and gave similar results; therefore, we report only on the larger set of maternal ratings.

### Genotype Data

All participants were genotyped on the Affymetrix 6.0 platform, which contains more than 900,000 SNPs. Quality control and imputation were performed on a larger dataset (N = 14,003) that also included genotype data from the parents of the twins. SNP data were cleaned with the following criteria: Hardy–Weinberg equilibrium (HWE) $p$ value >.00001, minor allele frequency (MAF) >0.01, call rate >0.95, concordance rate in duplicate samples >0.98, Mendelian error rate <0.02, and allele frequency difference with reference set <0.20. C/G and A/T SNPs