

8. ● Summary & general discussion

The aim of this thesis was to identify genetic and environmental risk factors for behavioral problems, in particular Attention Problems (AP) and Attention Deficit Hyperactivity Disorder (ADHD), and to quantify the contributions of latent and measured genotypes and environmental factors to the observed phenotypic variation. Longitudinal information collected by the Netherlands Twin Register (NTR) over the past 25 years was used to investigate early predictors of AP in childhood. In addition, information from DNA markers available in a subsample of participants was combined with the behavioral data in genetic association studies. Alongside this work, genome-wide association (GWA) studies for other phenotypes in childhood were performed for the EARly Genetics & Lifecourse Epidemiology (EAGLE) and Early Growth Genetics (EGG) consortium; these are two large international consortia of population-based cohorts that perform genetic association studies on a range of childhood phenotypes, including behavioral and somatic traits. Within the EAGLE consortium, I performed a meta-analysis on continuous measures of ADHD symptoms. Data from the Psychiatric Genomics Consortium (PGC),¹ a consortium that focuses on large GWA studies of clinical psychiatric cases and controls including childhood ADHD, were used to test whether the association results based on a study of clinical ADHD were predictive of AP in the NTR. Below I present a summary of the main findings of each chapter included in this thesis; next, I will discuss the results and offer a perspective on future research.

SUMMARY

In chapter two I looked at two early indicators of childhood temperament, i.e. items on “crying without a cause” and “being easily upset”. These items were included in a questionnaire sent out to parents of two-year old twins. The heritability of crying without a cause was estimated at 60% in boys and girls; the heritability of being easily upset was estimated at 43% in boys and 31% in girls. The shared environment explained 35 to 63% of the variance. Exploration of the large contribution of the shared environment indicated that it was not accounted for by the influence of birth cohort, gestational age, socioeconomic status, parental age, parental smoking behavior and alcohol use during pregnancy, or personality characteristics of the mother. The association between the two items (polychoric correlation of 0.36) was explained both by genetic and shared environmental factors. Importantly, both items were predictive of internalizing problems, externalizing problems and attention problems 5 years later, at age seven, with effect sizes between 0.28 - 0.42.

Chapter three reports on a study on the effect of birth weight on AP in childhood. In a sample of > 29,000 children, a lower birth weight was associated with higher AP scores at age 3, 7, 10 and 12 years. In monozygotic and dizygotic twins discordant for birth weight, the twin with the lower birth weight scored significantly higher on the continuous AP

scale. The differences in AP scores that were observed within unrelated pairs of children and within monozygotic and dizygotic twin pairs were of similar size. This finding strongly supports a causal effect of birth weight on AP. Twin pairs concordant for low birth weight but discordant for catch-up growth showed similar AP scores; the observed relationship between birth weight and AP was therefore unlikely to be explained by catch-up growth in low birth weight children.

Chapter four describes an attempt to replicate a previously reported interaction effect between a measured environmental exposure, namely breastfeeding, and measured genetic polymorphisms of the fatty acid desaturase 2 (FADS2) gene on IQ. These analyses were extended to include overactive behavior at age 3, AP at age 7, 10 and 12, and educational attainment at age 12 as outcome variables in addition to IQ. After correction for maternal education, a small effect of breastfeeding was observed for educational attainment at age 12, overactive behavior at age 3 and IQ across age 5-18; the latter effect was only marginally significant ($p=0.05$). The polymorphisms in the FADS2 gene showed no main effect on any of the included phenotypes; neither did they moderate the effect of breastfeeding.

Chapter five reports on an analysis in which AP in children who take part in NTR research projects were predicted based on polygenic risk scores for ADHD from an independent discovery sample. The results of the most recent ADHD GWA meta-analysis on 5,621 clinical ADHD cases and 13,589 controls, which was carried out by the PGC, were used to calculate polygenic risk scores in Dutch twins and predict their continuous AP scores. Polygenic scores were obtained by multiplying the number of observed effect alleles by the effect size found in the meta-analysis, summed over all loci (a locus refers to a location on the genome where a Single Nucleotide Polymorphism (SNP) was assessed). In the population based NTR sample, the ADHD polygenic risk scores were predictive of preschool and school-age maternal ratings and school-age teacher ratings of AP.

Chapter six includes the first report of a GWA meta-analysis of continuous measures of AP and ADHD symptoms that was conducted within the EAGLE consortium. Results of nine population-based cohorts were included leading to a total sample size of 17,560 children with genotype data imputed against the 1000 Genomes reference set. The Manhattan plot showed promising signals, but as expected with this sample size, no genome-wide significant findings were found at the stringent p-value threshold of $5E-8$. The tests of individual genetic variants were supplemented with gene-based tests and pathway analyses, but no signals were significant at a false discovery rate of 5%.

Chapter seven documents the results of several more GWA studies of behavioral and somatic phenotypes that were run for the EAGLE and EGG consortia, including atopic

dermatitis, height in childhood and puberty, body mass index (BMI) in childhood, pubertal staging and motor development. The NTR results have been uploaded for meta-analysis, together with the results from other childhood cohorts from Australia, Finland, the UK, Germany and other Dutch cohorts like Generation R. Genome-wide significant findings have already been published for eczema and pubertal height growth, and reported from preliminary meta-analyses of pubertal staging and BMI in childhood.^{2, 3} This chapter contains a detailed description of the genotyping, data cleaning and imputation procedures that preceded the genetic association analyses.

DISCUSSION

Although the identification of genetic risk factors for psychiatric disorders turns out to be much more involved than initially anticipated, one could argue that an even bigger challenge is to identify the environmental risk factors for psychiatric disorders in children and adults. One design that lends itself for the ‘environmental enterprise’ is the cotwin control design, which controls for a wide range of confounding factors, thereby eliminating important non-causal explanations for observed associations between environmental exposure and outcome measures such as AP and ADHD. The confounding factors that are controlled for in this design include genetic factors and environmental factors shared by twins like smoking during pregnancy and gestational age. The NTR data confirmed the association between low birth weight and AP, a finding that was also reported in two previous studies of twins discordant for birth weight. Together, these studies provide strong evidence for a causal relationship between birth weight and ADHD symptoms and underscore the importance of intervention programs that aim to prevent low birth weight.^{4,5} Moreover, they indicate that a close survey of these children at school age may be advisable; AP is negatively correlated with executive functioning, IQ and educational attainment in childhood and there is evidence that low birth weight children with learning disabilities are often not receiving the special academic assistance they need.^{6,7}

A low birth weight is one of the first measurable phenotypes in children. NTR characterizes the early development and environment of twins and multiples by asking mothers about early temperament and breastfeeding. Breastfeeding has been suggested to be beneficial for cognitive development; I looked at the association between breastfeeding on the one hand and IQ, educational attainment and overactive behavior at age 3 on the other hand. In these types of association analyses, maternal IQ and education can confound the association between breastfeeding and these outcomes, as mothers with a higher education are more likely to breastfeed their offspring.^{8,9} Given the high concordance of breastfeeding within twin pairs, it was not possible to apply the cotwin control design to the association of breastfeeding and cognition, but we controlled for maternal education in these analyses, as this information is collected in NTR surveys. The results presented

in chapter three and four thus underscore both the importance and difficulty to address issues of confounding.

"No aspect of human behavior genetics has caused more confusion and generated more obscurantism than the analysis and interpretation of the various types of non-additivity and non-independence of gene and environmental action and interaction..." (Eaves et al., 1977).¹⁰ This statement seems as true today as when it was written. I looked at a previously reported gene-environment interaction (GxE) effect between breastfeeding and variants in the *FADS2* gene on IQ and did not find evidence for this specific gene-environment interaction; an Australian group earlier reported another non-replication.¹¹ The only other replication effort performed before found a significant interaction, but in the opposite direction as found in the initial study.¹² These results are representative of the literature on candidate gene association and candidate GxE studies, as the field has been plagued by an inability to detect replicable results. This phenomenon may be due to difficulties to select appropriate candidate genes based on currently limited biological knowledge, and the small effects of individual genetic variants that we have now learned from GWA studies. In the context of small effects, small sample sizes and a publication bias towards positive results, false-positive findings are likely to occur.¹³ This issue is further complicated by the fact that interaction effects can be mere statistical phenomena, depending fully on the chosen model and scale of measurement, or anomalies in the data distribution; in these cases non-replication can be due to subtle differences in methodology and data distributions.¹⁴⁻¹⁶ In a critical review of candidate gene GxE findings in psychiatry,¹⁷ many more novel GxE studies than replication attempts turned out to be statistically significant, suggesting that GxE hypotheses appear more robust than they actually are. Given the progress in gene-findings studies since the advance of genome-wide methods, these hypothesis-free approaches should also be applied in GxE research. Finally, another explanation for the lack of robust findings in the gene-environment literature that needs to be considered here, is that gene-environment interactions explain little of the variation in psychopathology. This might not be a fashionable view, but given the fact that little empirical data support the importance of gene-environment interactions so far, this is a possibility one needs to consider seriously.

This thesis includes the first large-scale genome-wide association (GWA) study on continuous measures of Attention Problems and ADHD symptoms based on data in 17,560 children from nine population-based cohorts from the Netherlands, Germany, the UK, Australia, Spain and Norway. Although this is a large sample size in absolute terms, it is still modest if we compare it to sample sizes of other consortia such as GIANT (the Genetic Investigation of ANthropometric Traits), MAGIC (the Meta-Analyses of Glucose and Insulin-related traits Consortium) and SSGAC (Social Science Genetic Association Consortium) which meta-analyze data in up to 250,000 participants. I carried out a power

analysis for AP/ADHD, assuming a normal distribution of the phenotype and no phenotypic heterogeneity. Results showed that we had 80% power to detect a genetic variant that explained 0.23% of the variation (estimated with the Genetic Power Calculator).¹⁸ As we did not detect genetic variants at genome-wide significant levels, it must be assumed that effect sizes are even smaller. The power in our analysis is comparable to the latest PGC meta-analysis of childhood ADHD that included 5,621 clinical ADHD cases and 13,589 controls, and also did not detect genome-wide significant variants.¹⁹ For ADHD, and for other psychiatric phenotypes such as adult major depressive disorder (MDD), it has proven difficult to detect genetic variants in GWA studies. The latest PGC GWA study on major depression found no genome-wide significant hits when including 17,929 cases and 34,693 controls.²⁰ This is in contrast to the successes that have been reported for schizophrenia in particular, but also bipolar disorder. 108 Loci have been reported in the latest PGC analysis on schizophrenia that included 35,476 cases and 46,839 controls.²¹ For bipolar disorder, 8 genome-wide significant hits were reported in a GWA that included 13,741 cases and 19,762 controls.²² These successes were achieved in samples with much larger statistical power than available for ADHD and MDD; it should be kept in mind that the statistical power depends not only on the sample size, but also on the proportion of cases and the prevalence of the disorder.²³ As the prevalence of schizophrenia and bipolar disorder is much lower than for MDD and ADHD, cases represent a more extreme phenotype, which enhances statistical power substantially. Sample sizes for pediatric samples are generally smaller, and genome-wide studies of SNPs have so far not led to genome-wide significant variants in studies of autism, ADHD, and internalizing behaviors,^{19, 24-26} although for e.g. autism findings of the involvement of copy number variants (CNVs) and rare variants have been reported. Similarly, for AP and ADHD, involvement of rare CNVs has been observed.^{27, 28}

When analyzing genetic variants that are relatively common, such as SNPs, it is of importance to establish that these genetic variants contribute to variance in ADHD symptoms. This can be done by using several different methods, including so called chip-based heritability and polygenic risk score analyses. These chip-based heritability analyses employ methods for the estimation of narrow-sense heritability that use measured genome-wide genetic variants in large groups of unrelated subjects, rather than employing the theoretical values of genetic resemblance in relatives. There are two chip-based heritability approaches that differ substantially, with one approach resembling the variance decomposition methods as used in twin studies,^{29, 30} and the other based on density estimation (DE) methods.³¹ The first method requires raw genotype data and uses these to obtain a measure of genetic similarity between all possible pairs of unrelated individuals in the study. In a second step, this genetic relatedness matrix (GRM) is used to predict the phenotype similarity between individuals. The DE method can be applied after a genome-wide association study has been done. Here, the distribution of z-statistics from a GWAS is compared to the theoretical null distribution of z-statistics representing no effects.

Two studies that used the GRM method to assess the chip-based heritability of ADHD and ADHD symptoms in childhood found contradicting results. Whereas the PGC estimated a chip-based heritability of 28% for the liability to ADHD, a study of self, parent and teacher ratings of ADHD symptoms as measured with the SDQ and Conners' rating scale in the Twins Early Development Study (TEDS) found chip-based heritability estimates close to zero.^{32,33} There is no obvious explanation for this discrepancy, especially in light of the fact that I found that polygenic risk scores based on the PGC ADHD meta-analysis predict AP scores in a population based cohort such as TEDS as well. It is also important to note that the study from TEDS included a wide range of psychopathology ratings, all with chip-based heritability estimates close to zero, in contradiction with other studies that found chip-based heritability estimates of 13-43% for internalizing problems and 18% for social communication traits.^{25, 34}

The second approach to test for the relevance of common SNPs makes use of the effect sizes obtained in GWA studies to calculate genetic risk scores based on measured genotypic information in an independent set of individuals. These polygenic scores can then be tested for their ability to predict the phenotype studied in the GWA; if the prediction is significant, this indicates that the effect estimates from the GWA in aggregate contain a relevant signal for the disease or trait. With regard to ADHD, ADHD polygenic risk scores have been shown to be predictive of ADHD case status, and, as described in chapter five of this thesis, of measures of Attention Problems in a population-based sample.^{35, 36} The latter hints to the possibility to use polygenic scores to test for the genetic overlap across traits, in which polygenic risk scores based on one trait are used to predict the other trait in an independent sample. Such a cross-trait analysis is also possible with the GRM method, as it allows for the estimation of a chip-based co-heritability of two traits. Thereby, these methods contribute to the exciting shift from the analysis of single traits to cross-disorder studies. The PGC group, for example, has performed a GWA of 33,332 cases of five psychiatric disorders (autism spectrum disorder, attention deficit hyperactivity disorder, bipolar disorder, major depressive disorder, and schizophrenia) and 27,888 controls and detected four loci that contribute to variation across these disorders.³⁷ Moreover, using both polygenic risk score and chip-based co-heritability methods, they found convincing evidence for a genetic overlap across several psychiatric disorders, most convincingly bipolar disorder and schizophrenia.^{32, 37} These studies thus show that even when individual genetic variants confer only small relative risks, analyses that consider all risk variants together can contribute to an enhanced understanding of the etiology of psychiatric traits. In addition, they can be used to investigate the heterogeneity of psychiatric disorders and the genetic associations with endophenotypes, such as cognitive functioning.

Before discussing how research on AP and ADHD should proceed, I will also discuss the findings of this thesis in the light of the ongoing debate, both in the literature and in

the popular media, on the validity of ADHD as a clinical psychiatric disorder. Opponents claim that ADHD is a social construct leading to unnecessary medicalization of difficult children. They argue that there is no clear physical cause of ADHD and that the steep increases in ADHD diagnoses and stimulant use over time imply an important influence of social factors. Others defend the ADHD concept by pointing out that the observed behavior constitutes an impairment to the child and is associated with long-term negative consequences.³⁸ Although large heritability estimates and reported genetic associations are sometimes taken as evidence for the reality of ADHD as a disorder, it is important to note that the question whether ADHD is a 'real' medical condition or a social construct cannot be answered by looking only at genetic effects. Many so-called social constructs such as education or socioeconomic status turn out to be influenced by genes, while some somatic diseases such as lung cancer clearly have a huge non-genetic component to explain their etiology.

Yet, genetic analyses are of importance to aid our understanding of ADHD. The polygenic risk score analysis described in chapter five indicates that the genetic liability to ADHD is continuously distributed. This implies that the clinical cutoff applied in clinical practice is in essence arbitrary; one's score on the ADHD scale can be more or less extreme, with the subsequent impairment also partly depending on the context in which you are functioning. This idea is further supported by the fact that the number of reported AP and ADHD symptoms has stayed constant over the last two decades, while the number of ADHD diagnoses and stimulant prescriptions has seen a steep increase in this period, as is indeed also reported by parents of NTR twins, thereby signifying that which score on the continuous scale is considered impairing is partly influenced by social factors. Moreover, it is of interest to note that children in the youngest one third of a class have a 50% increased risk of being prescribed stimulant medication when compared to the oldest one third. This is probably due to the fact that the same behavior is considered more impairing when children are surrounded by somewhat older children that have matured to higher levels of self-control.³⁹

The notion that children with a certain level of ADHD symptoms are now more often considered clinical cases than before, does not answer the question whether this is because the clinical disorder is now recognized more often, or whether the increase is due to misclassification, or whether the number of symptoms stayed the same while the environment of the child became less tolerant of the behaviors. In the end, questions on the validity of the ADHD diagnosis can only be answered by a careful reflection on the nature of psychiatric disorders, the borders between psychiatry and normality and the difference between treatment and enhancement; a clear-cut answer cannot be provided. Given the earlier mentioned negative long-term consequences and impairment associated with extreme scores in current society, there is no reason to put ADHD aside

as a nonsense diagnosis; the impairment is a reality that needs to be reacted upon. Yet, the understanding of ADHD as the extreme end of a continuous distribution can hopefully help to lessen stigmatization and to keep an open view on the possibilities of treatment for individuals at threshold and subthreshold levels of the disorder.

FUTURE PROSPECTS

If we accept that there is convincing evidence for the implication of genetic variants in the etiology of psychiatric traits and ADHD, it is clear that GWA studies will continue to play an important role in genetic studies of ADHD in the near future. Such studies will enlighten us about the genetic architecture of traits and will, maybe paradoxically, also aid research on environmental factors by enabling methods that go beyond establishing an association between the environment and a trait, to explain the underlying cause of the association.

An important question is how GWA studies should move forward in order to find replicable associations. Clearly, increasing sample sizes is of major importance for success; current efforts to collect GWA data in larger sets of ADHD cases are thus worthwhile. Several issues need to be considered when one aims to increase sample sizes for GWA studies. It is for example, of importance to look at the implications of the organizational structure of GWA consortia. Two different strategies are currently followed; the PGC is the most prominent example of a consortium that requires all participating cohorts to upload their individual-level genotype and phenotype data to a common server. The main advantage of this approach is that quality control, imputation and analyses can be performed on the total dataset, enhancing uniformity across cohorts. Once the data have been uploaded, the speed of the project is dependent on the consortium analysts only; in general, this will speed up the process. Moreover, this approach allows for more complex analyses for which individual level data are required, such as the GRM chip-based heritability estimates. A clear downside of this approach is that the requirement of individual level data sharing hampers the number of cohorts that are able or willing to participate. Sharing of individual level data comes with many ethical and legal issues, so that some cohorts will not be able to contribute. Moreover, it is not easy to find a balanced system to distribute academic credit over the contributors; once data have been uploaded, contributing cohorts are in a somewhat difficult position to negotiate. To bypass these problems, other consortia only require the sharing of GWA summary statistics, an approach that is taken in for example ENIGMA, MAGIC and GIANT and also in the EAGLE and EGG consortia. To be effective, it is essential that there is a clear and comprehensive analysis plan and that cohorts are able to provide high quality results within a reasonable time frame. In practice, this strategy has proven to work well, both in the EAGLE and EGG consortia and in other consortia of even larger scale, such as the SSGAC that recently included 126,559 individuals from 54 cohorts in a report on three genome-wide significant SNPs for educational attainment.⁴⁰

Another important issue when it comes to increasing sample size is the issue of phenotype definition. For studies focusing on gene finding rather than precise effect estimation, a cost-effective approach may be to collect large samples with less in-depth, but easier to obtain phenotypic information rather than extensive clinical assessments; this type of phenotype measure is often collected in large twin registers. The disadvantage of survey measures may be compensated by the increase in sample size and availability of longitudinal phenotype information; using this information may be a promising approach for future projects. Another option is the use of electronic medical records and information on medication use as a proxy for case status, or the use of single questionnaire items that ask a participant whether he has ever been diagnosed with a particular disease. The latter can be a fruitful approach, as shown by the GWA on atopic dermatitis reported in chapter seven, and studies that successfully replicated known genetic associations with self-reported health data.^{3,41} However, it should be noted that the use of a single item phenotype definition will probably be more challenging for diseases with a low prevalence, such as for ADHD (the prevalence of ADHD is around 5%, whereas eczema has a prevalence around 15-30% in childhood).⁴² In these cases, the collection of behavior questionnaires that assess the full range of behaviors in the population remains an attractive alternative, as was done in the GWA meta-analysis on continuous measures of ADHD symptoms described in chapter six. These population-based cohorts have further advantages in terms of cost-effectiveness, as they generally assess a wide range of phenotypes, instead of being focused on one particular disease only. Another worthwhile option is to perform a meta-analysis in which data from clinical case-control and population-based cohorts are analyzed together. Although the heterogeneity in phenotypic assessment will lead to a decrease in statistical power, this effect could very well be outweighed by the increase in sample size. With regard to ADHD, it will be worthwhile to start with a combined meta-analysis of the PGC and EAGLE data. As an extension of such a cross-consortium meta-analysis, the inclusion of adult data could be considered, as behavior genetic studies have found a considerable overlap of genetic variants for ADHD symptoms across the life span.⁴³

Obviously, this approach is only fruitful if the genetic variants for a clinical diagnosis of ADHD overlap to a sufficient extent with genetic variants that influence ADHD symptoms in the general population. In line with previous behavior genetic studies based on twin data, our polygenic risk score analysis indicated a significant genetic overlap between the genetic variants detected by GWA of clinical ADHD cases and continuous measures of ADHD as measured with the CBCL AP scale in the NTR in childhood.^{44,45} As a follow-up on this analysis, I tested for the overlap between the GWA meta-analysis results of PGC and EAGLE; the latter included a considerably larger number of children and a wider variety of measures of ADHD symptoms in terms of measurement instrument, age and rater. Selecting independent top SNPs from the PGC analyses after appropriate quality control, the top hits of PGC did not show significant overall inflation of p-values in EAGLE (as

indicated by the inflation factor lambda). However, the direction of effect was significantly more consistent across the consortia than expected by chance, as indicated by a sign-test (see Table 1), with a stronger level of significance at less stringent p-value thresholds of inclusion. This consistency of effect direction at more liberal p-value thresholds shows that there is a relevant genetic signal in the PGC ADHD mega-analysis that overlaps with the genetic signal in the EAGLE meta-analysis of ADHD symptoms, providing support for the above-mentioned possibility to perform a cross-consortium meta-analysis of PGC and EAGLE. At the same time, the most significant top hits from PGC (p-value < 1E-5) showed a similar direction of effect in the EAGLE meta-analysis in only 50% of the cases; this again underscores the difficulty to identify individual genetic variants for ADHD with current sample sizes. It will therefore be worthwhile to not focus too extensively on the top hits, but instead search for replication for a longer list of variants, thereby incorporating also biological information in the process of prioritizing variants for follow-up analyses.

Table 1. Overall inflation and consistency of direction of effect in the EAGLE ADHD symptoms meta-analysis for the independent top SNPs from the PGC ADHD case-control mega-analysis.

SNP set, p-value threshold PGC results	n SNPs	Lambda	empirical p-value	Sign-test (same/different direction)	p-value sign-test
1E-5	12	1.813	0.1555	6/6	1
1E-4	75	0.808	0.7808	45/30	0.1053
1E-3	600	0.965	0.6621	322/278	0.7909
1E-2	4,563	0.992	0.6526	2,409/2,150	1.3E-4
1E-1	32,623	1.021	0.0766	17,045/15,562	2.2E-16
1	168,871	1.005	NA	85,735/82,747	3.4E-13

Different methods have been developed to analyze the genetic architecture of traits and their chip-based heritability. These methods have not yet been extensively tested for their robustness against violations of assumptions. Moreover, the effects of specific thresholds applied in quality control steps are often unknown. It is currently, for example, unclear how quality control filtering affects the estimates of explained variance from the GRM and DE methods, and the outcomes of polygenic risk score analyses. This includes filtering on MAF, imputation quality and HWE, but also correction for patterns of linkage disequilibrium. The latter is an important practical issue, obtaining a set of independent variants is essential for e.g. the DE method, and may be of importance for polygenic risk score analyses; but both the appropriate methods and cutoffs are yet to be agreed on. In studies that apply polygenic risk score analyses, results are often reported for both pruned and unpruned datasets, but whenever results are dissimilar, it is hard to judge which result is more appropriate. For the GRM method, several options have been proposed to correct for LD, but it remains to be established which method performs best under which scenario^{46, 47}

With regard to GWA studies, outstanding issues include variability in QC thresholds both at the individual cohort and meta-analysis level, the appropriate cutoff for genome-wide significance when analyzing 1000 Genomes imputed datasets and the use of fixed versus random effect models. Obviously, this situation is a consequence of the pace of development in the field, and is as such inescapable. Nevertheless, studies that look further into these issues are timely and needed. Where discrepancies have been found such as for the above mentioned differences in chip based heritability estimates of ADHD and ADHD symptom scores, these studies may help to elucidate whether these differences can be attributed to methodological differences. The last year has seen a positive development in this regard. Papers by Dudbridge et al. and Lee et al. on polygenic risk score analyses provided a helpful examination of the behavior of these tests under different scenarios, although they do not address the issue of pruning.^{48, 49} With regard to the GRM method, a recent paper has tested the effects of violations of five main assumptions.⁴⁷ More of these methodological issues will likely consolidate in the near future, increasing the validity of these methods and the consistency of their application.

Meanwhile, research on environmental risk factors should also proceed. Several methods have been suggested to test for causality of associations with environmental risk factors, including randomized controlled trials, statistical control for measured confounding factors, twin and family designs and Mendelian randomization.⁵⁰ While each method has its own limitations, together these methods can provide convincing support for causal effects of established risk factors. So far, they have been applied somewhat sparsely to environmental risk factors for ADHD. In addition to the studies I have performed, two other studies, one applying a discordant sibling design and another employing an assisted reproduction designs of mothers that are genetically unrelated to their child, suggest that the relationship between smoking during pregnancy with ADHD is at least partly due to confounding by genetic and other familial factors.^{51, 52} A sibling study found that the association between low income during early childhood and ADHD remained after control for familial factors.⁵³ Further use of these methods can hopefully contribute to the clarification of the role of the environment in ADHD and identify genuine risk factors.

The lack of significant findings in GWA studies is sometimes attributed to the hypothesized contribution of gene-environment or gene-gene interactions to the phenotypic variation. However, given the above-mentioned lack of empirical evidence for the contribution of gene-environment interaction to psychopathology, the initial focus of current studies should be on the identification of robust genetic associations. Subsequently, gene-environment or gene-gene interactions can be investigated. Given the growing number of established associations and the clear predictive value of polygenic scores, studies that combine GWA data with measured environmental factors are a clear way forward to the field.^{54, 55} One example of such a study is provided in chapter seven, where we

contributed to a genome-wide study on the interaction of smoking during pregnancy and genetic variants on birth weight. A similar argument can be made for studies on gene-gene interactions. Although the number of possible interactions is extremely large and the power for interaction studies is generally lower than for main effects, the multiple testing burden can be reduced substantially by including only interactions between known loci, a strategy that was used in the EAGLE GWA on atopic dermatitis, reported in chapter seven.

It is not expected that genetic variants will be suitable for risk prediction and prevention in the near future.⁵⁶ However, there have already been examples of pharmacological agents based on genetic findings.⁵⁷ Such examples show that the identification of genetic variants can be very valuable, even when they do not explain all heritability. In the end, the hope of many who work in psychiatric genetics may thus prove grounded; that psychiatric genetic studies can make an important contribution to the improvement of mental health.