Chapter 8
Summary & General discussion
SUMMARY OF THESIS FINDINGS

This thesis aimed to study the complex link between genetic effects and measured environmental risk factors in major depressive disorder (MDD) in empirical data, and to explore boundaries of the consequences of two Genome-Wide Association study (GWAS) designs and assortative mating from a theoretical perspective. Both genes and environment affect MDD risk, but it remains unclear whether both act independently, whether they interact, or whether environmental risk might actually reflect shared genetic effects between MDD risk and e.g. behavioral traits that might increase environmental stress.

The most studied candidate gene in MDD is the serotonin transporter gene. The length polymorphism in the promoter region of this gene (5-HTTLPR) has been hypothesized to influence MDD risk, because an important group of antidepressant drugs acts on the serotonin transporter. A Science paper by Caspi et al suggested that this gene mainly exerts its influence when persons had been exposed to childhood trauma. In Chapter 2, the 5-HTTLPR polymorphism was analyzed in 1593 cases and 1411 controls from the Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Twin Register (NTR). In these individuals, four different outcome measures were defined: lifetime MDD, suicidal MDD, chronic MDD, and course of MDD (chronic versus non-chronic). No evidence was found for either direct effects of 5-HTTLPR on these outcome measures or interaction effects between 5-HTTLPR and five environmental risk factors for MDD: lifetime stressful life-events, recent stressful life-events, sexual abuse, childhood trauma, and educational attainment (as proxy for social economic status associated with increased stress).

In Chapter 3, the relation between MDD and educational attainment was investigated in approximately 25,000 individuals from the Psychiatric Genomics Consortium (PGC wave 1) with additional Dutch and Estonian data. An increased risk for MDD was confirmed in individuals with lower educational attainment. Subsequently, the possible contribution of shared genetic effects to this link was assessed with three different methods applying data of 884,105 autosomal common single-nucleotide polymorphisms (SNPs). Firstly, polygenic risk scores (PRS) based on GWAS results on education attainment in ~120,000 individuals (EA-PRS) did not affect MDD risk, and PRS based on MDD GWAS results in ~20,000 individuals (MDD-PRS) did not affect EA. Secondly, a non-consistent weak significant negative genetic correlation was found with bivariate genomic-relationship-matrix restricted maximum likelihood (GREML). Thirdly, no concordance was found in either significance or direction of SNP effects across MDD GWAS and EA GWAS results based on SNP effect concordance analysis.
(SECA). To conclude, these findings indicate that it is unlikely that shared genetic effects explain a large proportion of the link between MDD risk and lower education attainment, but a small genetic contribution to this deleterious link could not be excluded.

Up to 2014, research on gene-by-environment (GxE) interaction in MDD had mainly focused on candidate genes, such as 5-HTTLPR. However, with the emergence of MDD cohorts with genome wide SNP data, novel methods were developed that allowed to tag genome-wide genetic MDD risk with polygenic risk scores (PRS). In Chapter 4, PRS were constructed in 1645 cases and 340 controls from NESDA based on discovery results from the large Psychiatric Genomics Consortium (PGC wave 1). These PRS were found to have an increased impact on MDD risk in individuals exposed to childhood trauma (CT), which suggested gene-by-environment interaction on a genome-wide scale. This interaction-effect was found both as departure from multiplicativity (combined impact of PRS and CT larger than the product of the individual effects) and as departure from additivity (combined impact larger than the sum of the individual effects), the latter of which has been hypothesized to be more plausible from a biological perspective.

In Chapter 5, the interaction between polygenic risk scores (PRS) and childhood trauma (CT) was further tested in seven cohorts from PGC (wave 2) with CT information available in 3,024 cases and 2,741 controls. CT had consistent impact across cohorts, with similar impact in males and females. However, the interaction effects were heterogeneous with a positive interaction effect in NESDA (as in Chapter 4), negative interaction effect in the Radiant-UK study, and no interaction in the other contributing five cohorts, resulting in no overall evidence for interaction between PRS and CT in MDD. The results from Chapter 5 illustrate the heterogeneity of MDD, and suggest that the results of Chapter 4 can best be interpreted as a single cohort phenomenon.

The focus was switched from analyzing empirical data to theoretic work in Chapter 6. Two GWAS study designs applied in the Psychiatric Genomics Consortium (PGC) were considered with respect to their power of SNP association analysis and SNP-heritability estimates (proportion of population-variance in disease-risk attributable to genome-wide common SNPs). First, parent-affected-offspring trio data are regularly applied in the subgroups of the PGC analyzing autism spectrum disorder (ASD) and attention deficit/hyperactivity disorder (ADHD). Trio data are essential to detect de novo mutations, but its use may results in reduced power in association analysis and underestimation of the SNP-heritability compared to analyses in case screened-control data. This difference is attributable to (i) potential oversampling of multiplex families (with more than
one affected offspring) and to (ii) assortative mating, which describes the correlation between mating partners in vulnerability for ADHD or ASD. Second, the use of poorly or unscreened controls for common disorders will result in decreased power in association analysis and decreased SNP-heritability estimates. In particular, for MDD with a lifetime prevalence of around 15%, the anticipated underestimation of the SNP-heritability was analytically derived at 28% when none of the controls would be screened. An updated equation was provided to properly scale the SNP-heritability when including unscreened controls (Equation 3 in Chapter 6). When aiming to analyze the polygenic effects in psychiatric disorders, it is advisable not to use trio data and to properly scale the SNP-heritability when applying data with unscreened controls.

Research has found significant partner-resemblances for psychiatric disorders, that is, mating partners are more often concordant in psychiatric disorder-status than expected by chance. This phenomenon is often referred to as assortative mating. However, the potential consequences of these partner-resemblances have not been quantified and have been left implicit despite available theory in the quantitative genetics literature. Therefore, in Chapter 7, boundaries were quantified for the anticipated consequences for disorder prevalence and heritability under various inevitable assumptions. The consequences are most pronounced when partner-resemblance is attributable to phenotypic assortment (partner-resemblance driven by the psychiatric trait), and are reflected in increased population prevalence and heritability in the offspring generation. From the first generation in which assortative mating takes place, the consequences add generation after generation to reach equilibrium asymptotically over generations. Because of this equilibrium, assortative mating is unlikely to balance the impact of reduced fecundity of psychiatric patients in the long term, as analytically derived in Chapter 7. Modeling suggests that the heritability of none of the psychiatric disorders considered is likely to increase with more than 5% from one generation of assortative mating (or 13% over several generations). The population prevalence will increase most for rare disorders with high heritability, such as the prevalence of ASD that might maximally increase 1.5-fold after one generation of assortative mating (or 2.4-fold over several generations). While emphasizing the limitations inherent to the inevitable model assumptions, genetic theory suggests that the consequences of assortative mating are, at most, modest for the heritability, but may be considerable for the population prevalence.
DISCUSSION OF THESIS FINDINGS

GxE interaction research with candidate genes

Over a decade of research on gene-by-environment (GxE) interaction with candidate genes in major depressive disorder (MDD) has led to contradicting findings from which a pattern of non-replication has emerged. Chapter 2 of this thesis further strengthens the pattern of non-replication of the interaction effect between 5-HTTLPR and childhood trauma presented by Caspi et al.1 In addition, a critical review from Duncan et al suggested that the GxE literature from the candidate-gene era suffered from publication bias, because 96% of novel GxE studies yielded significant results compared to only 27% of replication studies, and because smaller replication studies reported more significant results than the larger samples.2 To conclude, it seems unlikely that the Caspi finding is generalizable to other cohorts, and there doesn’t appear to be much ground to further study interaction between childhood trauma and 5-HTTLPR in MDD.

GxE with genome-wide information

Chapter 4 of this thesis describes a positive statistical interaction between polygenic risk scores and childhood trauma in the Netherlands Study of Depression and Anxiety.3 In the context of the critical review of Duncan et al of the literature on GxE with candidate genes,2 was it worth to publish this novel finding while based on one cohort only? I think it was, as these analyses had several marked differences with the candidate gene literature. First, polygenic risk scores have a significant impact on MDD that is consistent across different cohorts,4-6 which contrasts the small impact of 5-HTTLPR on MDD that was found in a large meta-analyses showing considerable variety across individual cohorts.7 Interaction effects can result in only a small main genetic effect, but this requires that the effects in different environmental strata balance exactly. Testing for interaction with polygenic risk scores with confirmed consistent main effects might be more powerful than testing for interaction with candidate genes with small inconsistent effects or no main effect at all.8 Second, although there were several different polygenic risk scores that could have been tested, for example based on a discovery GWAS on schizophrenia instead of MDD, I felt there was one candidate of most interest: the genome-wide set of polygenic risk scores based on the largest GWAS for MDD at the time.4 This apparent choice contrasted the many independent candidate genes that were tested contributing to multiple testing and suggested publication bias in the candidate gene literature.2 Nevertheless, there were of course many environmental risk factors that we could have tested. Exposure to childhood trauma was chosen as environmental
risk factor, because this is one of the strongest and most consistent risk factors with a lifelong impact on MDD risk, which has even been hypothesized to distinguish a neurobiological distinct subtype of MDD. Taken all together, I think that testing for interaction between polygenic risk scores and childhood trauma in MDD was not only novel, but also an obvious step to take at the time, and I think the NESDA finding was justifiably published as a single cohort result. Nevertheless, this single cohort finding confirming polygenic GxE interaction was, subsequently, not replicated by Mullins et al who found evidence for an opposing interaction effect. Analyses combining data of seven cohorts totaling 5,765 individuals showed no overall evidence for interaction in either direction (Chapter 5 of this thesis).

Interpreting PRSxCT results
How to interpret the positive interaction effect between polygenic risk scores (PRS) and childhood trauma found in Chapter 4 with NESDA data, contrasting the negative interaction effect found by Mullins et al with the Radiant-UK data? Most importantly, these contrasting findings seem to illustrate the genetic heterogeneity in MDD, which had already been indicated by Lee et al who found that the genetic coheritability between different MDD cohorts showed more variety than between e.g. different schizophrenia cohorts. In Chapter 5, an attempt was made to aid further interpretation with a simulation study. Therefore, several different scenarios of underlying genetic architecture of MDD and childhood trauma were simulated, followed by comparison of interaction-effect estimates with results from empirical data, and by comparison of the patterns of mean PRS in exposed controls, unexposed controls, exposed cases and unexposed cases. From simulation, the typical pattern seemed that exposed cases had lower PRS than unexposed cases, and exposed controls had lower PRS than unexposed controls, explained by the fact that those exposed to childhood trauma require less polygenic risk to become affected. Notably, this pattern does not reflect GE-correlation. In the full population, exposed controls are relatively rare as are unexposed cases. Thus, the mean PRS in all unexposed individuals shifts towards the mean in controls and the mean PRS in exposed individuals shifts towards the mean in cases, overall resulting in equal mean PRS in exposed and unexposed individuals in the full population.

It appeared that the NESDA-findings were unlikely to primarily represent different directions of SNP effects in exposed and unexposed individuals, because this would have resulted in a negative interaction-effect attributable to the discovery GWAS that would have primarily tagged effects in unexposed
individuals (the prevalence of CT is approximately 0.25). Rather, the NESDA-findings seemed to best fit the simulated scenario with either decreased environmental variance or increased genetic variance in exposed individuals, explaining the increased effect of PRS in exposed individuals. At first sight, the negative interaction effect in the Radiant-UK study appeared consistent with a different direction of genetic effects in exposed compared to unexposed individuals as just discussed. However, the mean PRS in Radiant showed a distinct pattern with larger PRS in exposed cases than unexposed cases, and larger PRS in exposed controls than in unexposed controls. This pattern seemed to fit simulated data with genetic effects impacting on CT (not ruling out difference of genetic effects, of course). However, this comparison between simulated and empirical data only provides a rough feeling for the genetic architectures that the NESDA-findings and Radiant-findings might be compatible with, because simulation didn’t allow for different genetic architectures between discovery and target set, and because simulation was based on several inevitable and partly arbitrary assumptions. In addition, we have to keep in mind that, although interesting from a theoretical perspective, the heterogeneity across the six cohorts studied in Chapter 5 does not justify too detailed interpretation. Nevertheless, I think more insight is obtained by attempting to understand patterns from PRS interaction-analyses with simulation, than when not attempting at all. Importantly, non-consistent findings in PRSxCT interaction in MDD do not suggest that no SNPxCT interaction effects exist in MDD. For example, if a proportion of 10% of MDD SNPs would be moderated by childhood trauma, their interaction-effects would likely be diluted in PRS-analyses, while interaction in 10% of affective SNPs would be a relevant phenomenon to study.

Methodological challenges in GxE interaction analyses
Several papers have described methodological challenges of GxE analyses in twin studies and candidate gene studies, and these challenges may also apply to tests for GxE analyses with polygenic risk scores. Purcell discussed gene-environment (GE) correlation, which describes the impact of genetic variants on environmental exposure via e.g. personality characteristics, or when considering childhood trauma, via a link of transmitted genetic variants with personality characteristics in the parents. Purcell explained that increased genetic effects in an environmental condition could result from moderated genetic effects (GxE), but also from risk variants being more likely to be present in that environmental condition as a consequence of GE-correlation. Notably, within NESDA and the Radiant-UK study, the PRS based on MDD discovery results were correlated to CT.
Therefore, a simulation study was performed in Chapter 5 to assess the potential impact of GE-correlation. These simulations indeed confirmed an inflated type I error rate in the context of GE-correlation, but to a modest extent of 0.075 (with alpha set at 0.05) for a strong correlation of 0.3 between G and E. This indicates that the NESDA and Radiant-UK findings are most likely to represent non-spurious single cohort phenomena.

Apart from GE-correlation, Eaves (2005) showed with simulation that even when ruling out gene-environment correlation, spurious GxE results can still be found when e.g. the disorder liability would be non-normally distributed, thereby reemphasizing the fragile nature of tests for interaction.13 In his work, Eaves did not provide means to disentangle spurious results from true interaction effects, but he predicted that interaction effects would be common rather than specific when they would be attributable to a non-normally distributed disorder liability. This pattern was not found in the years following 2005, although a considerable number of GxE studies were conducted.2 In 2014, Keller showed that many GxE studies might have overestimated or underestimated interaction effects by improperly correcting for covariates. That is, covariates were typically included only for their main effects whereas their interaction effects ought to also be included (covariates times G and covariates times E).14 The analyses in Chapter 4 did not comply with these recommendations, but in Chapter 5 the analyses were corrected also for the interaction effects of the covariates and showed a similar positive interaction effect within NESDA.

**Future perspective on GxE interaction research in MDD**

The non-consistent findings in PRSxCT interaction in MDD do not suggest that no SNPxCT interaction effects may exist in MDD. However, the question is how to best test for SNPxCT effects in MDD. The risk for publication bias with candidate genes (or SNPs) underlines the importance of a hypothesis free approach.2 Assuming that interaction effects, if existing, will be of the same magnitude as main SNP effects, GWAS samples with CT information of tens of thousands of cases and controls will likely be needed to reach genome-wide significance. Alternatively, a two-step approach could be applied to increase power by first selecting SNPs with e.g. a correlation with CT in the combined case-control sample and second testing these SNPs for interaction with CT in predicting MDD.15 The increase in power from a two-step approach can be considerably,15 and I think these analyses should be conducted within the context of MDD, although I wouldn’t be surprised if no consistent interaction effects would be detected given the large genetic heterogeneity of MDD.11
Assessing CT information in additional numbers of MDD cases and controls will assist SNPxCT research, but also PRSxCT analyses. Notably, the PRS applied in Chapter 4 and Chapter 5 were based on discovery GWAS results from samples with unknown mixtures of individuals exposed and unexposed to CT. This complicates interpretation of results and reduces power of interaction analyses. It would be preferable to have two distinct discovery GWASs, one exclusively based on unexposed individual and one exclusively based on exposed individuals. Unfortunately, current sample sizes did not allow for this approach, but these analyses might be of particular interest. There is a third possible advantage of assessing CT information in large numbers of cases and controls: a GWAS in unexposed individuals (approximately 75% of the original sample depending on the definition of CT) may provide increased power to detect SNPs associated to MDD, because unexposed individuals require more genetic risk on average to become affected.

An important challenge will come from the choice of the environmental factor to test, a choice for which no hypothesis-free work-around is available. I have argued that CT is a plausible candidate, because of its strong and lifelong impact on MDD risk, and because exposure to CT has been hypothesized to distinguish a neurobiological distinct subgroup of MDD. However, many other environmental conditions can also be tested, such as stressful life-events later in life, socioeconomic status or air pollution, inevitably introducing (hidden) multiple-testing burden risking false-positive findings. One might argue to adjust the level of significance according to the number of environmental factors tested, but this seems an unfeasible approach, in particular when tests for different environmental factors are presented in different papers. Indeed, a more plausible approach would be to emphasize the importance of independent replication, and to regard single-study results as no more than hypothesis forming.

In addition to the challenges with respect to power and choice of environmental risk factor, research on GxE interaction in MDD is further complicated and confused by the methodological challenges intrinsic to GxE testing. A choice needs to be made whether to test for interaction as departure from multiplicativity or as departure from additivity. Although the latter has been hypothesized to be more in line with meaningful biological interpretation, most interaction analyses test for interaction as departure from multiplicativity as these readily follow from logistic regression. The challenges and potential methodological pitfalls do not create optimism to test for GxE-interaction in MDD. However, when a well-replicated GxE interaction effect would be found, with well-understood biological interpretation with respect to e.g. gene-
inactivation by methylation, this would be of major importance for understanding the complex pathophysiology of MDD. Taken all together, I think two-step SNPxE analyses should be conducted in the short term, but I fear that the current MDD data may present too many challenges to properly test for SNPxE interaction. Therefore, I recommend that researchers who collect samples for GWAS studies consider to collect environmental information in a uniform manner to prepare for solid SNPxE testing on a large scale in the years to come. Alternatively, I would advice researcher to at least facilitate follow-up of study participants and obtain permission for record linkage.

**Depression, educational attainment, and genetic correlation**

In Chapter 3, data of approximately 25,000 individuals were applied to test whether the phenotypic link between lower educational attainment (EA) and increased MDD risk could be partly attributable to shared genetic effects (or LD between effective loci on MDD and EA). Therefore, three methods were applied: bivariate genomic relatedness matrix restricted maximum likelihood (GREML) analyses, EA polygenic risk scores predicting MDD, MDD polygenic risk score predicting EA, and SNP effect concordance analysis (SECA). None of these methods showed consistent evidence for genetic correlation between MDD and EA, indicating that genetic effects are not expected to explain much of the phenotypic link between MDD and EA. Notably, these analyses required considerable computational time, in particular the GREML analyses on the genomic relatedness matrix including 312,512,500 elements. Since then, a new method, LD score regression, has been developed illustrating the great progress in genetic research, because tests for genetic correlation can now be conducted within one minute with LD score regression based on GWAS summary statistics only. Indeed, although the analyses described in Chapter 3 resulted in a paper on its itself in 2014, the first LD score regression paper in 2015 presented the genetic correlations between 24 traits, i.e. 276 times the number of genetic correlation estimates. This study also found no evidence for genetic correlation between years of education and MDD risk. Interestingly, the Social Science Genetics Consortium recently found a negative genetic correlation between neuroticism and educational attainment of -0.42 (SE=0.07, p=2.8e-8), which may seem to contrast the finding of Chapter 3. This difference could be attributable to the more accurate estimation of genetic effects on educational attainment as sample size has increased from approximately 125,000 to approximately 300,000, and because MDD and neuroticism are closely linked but not quite the same with genetic correlation estimated at 0.66 (SE=0.09,
LD score regression has considerably narrowed the need for bivariate GREML and PRS analyses when aiming to test for the genetic correlation between two traits. LD score regression has slightly less power than GREML analyses and is not feasible when GWAS results are based on a small sample with less than around 5,000 individuals. For samples including less than 5,000 individuals, bivariate GREML and PRS analyses can be considered, given that individual level genotype data are available. For samples with less than around 3,000 to 4,000 individuals, bivariate GREML may be underpowered, and tests for genetic correlation can then be performed with PRS-analysis, given that adequately powered discovery GWAS results are available for one of the traits. Notably, in Chapter 3, the phenotypic link between lower EA and MDD risk was confirmed with phenotypic data, the analyses of which cannot be replaced with LD score regression.

Caution is required when testing for genetic correlation with bivariate GREML for disease-trait pairs, because of concerns raised by Golan et al about univariate GREML in estimating SNP-heritability in cases-control data. In case-control data, oversampling of cases introduces correlation between genetic and environmental effects, which violates the assumptions underlying GREML analysis and results in underestimation of the SNP-heritability. Golan et al advised applying cross-product Haseman-Elston (HE) regression, which is anticipated to yield unbiased estimates of the SNP-heritability. Unfortunately, Golan et al did not address bivariate GREML analyses in the context of case-control data, and I am not aware of any clear advice in this aspect. However, I would advice caution, as it seems plausible that case-control data may also result in biased estimates from bivariate GREML analyses. The proposed method of Golan et al to correct for covariates does not naturally extend to bivariate analysis. Here, I would advice a pragmatic approach, which is to present results from both bivariate GREML analyses and bivariate HE-regression with the residuals of regression of the two traits of interest on the relevant covariates.

**SNP heritability of psychiatric traits**

The proportion of variance in disease risk attributable to genotyped SNPs (SNP heritability) can be assessed with a variety of methods on different types of case-control data. The methods most often applied are genomic relatedness matrix restricted maximum likelihood (GREML), Haseman-Elston (HE) regression, LD score regression, and, less frequently, the method from So et al. SNP
heritability estimates are typically lower than family-study based heritability estimates, a phenomenon referred to as the missing heritability, which has been hypothesized to be attributable to, for example, non-genotyped genetic risk variants. Indeed recently, Yang et al. found negligible missing heritability for height and BMI when imputing rare genetic variants and applying GREML stratified for minor allele frequency and linkage disequilibrium. For case-control data, a different method has been proposed that constructs haplotypes from common SNPs to tag rare variants followed by HE-regression stratified by minor allele frequency. With this method a haplotype heritability of 0.64 for schizophrenia was found compared to SNP heritability of 0.32. In particular, SNP heritability estimates from case-control data also depend on the method applied as Golan et al found that GREML likely underestimates the SNP heritability, because of correlation between genetic and environmental effects introduced by oversampling of cases typical in case-control studies. Golan et al. found that HE regression is robust against oversampling of cases, and it can be assumed that the same holds for LD score regression, which provides similar SNP heritability estimates as HE regression.

In Chapter 6, the consequences of study design on SNP heritability estimates were explored, with a specific focus on the designs applied in the Psychiatric Genomics Consortium. As shown, the SNP heritability is likely underestimated when analyzing parent-offspring trio data, that is regularly applied in childhood onset disorders such as autism (ASD) and attention deficit/hyperactivity disorders (ADHD). In addition, the SNP heritability will be underestimated when analyzing data of unscreened or poorly screened controls for common disorders, such as major depressive disorder. An updated equation is provided for estimating SNP heritability from data with unscreened controls (Equation 3 in Chapter 6). This equation corrects the expected underestimation of $(1 - K \mu)^2$ for a disorder with a population prevalence $K$, that is studied in a sample with a proportion of unscreened controls $\mu$. Ideally, the SNP heritability underestimation from trio data would have also been captured in an equation, but this was considerably more complex theoretically, and found to depend on (the often unknown) proportion of multiplex families in the study, and the amount of assortative mating for the disorder under consideration. Nevertheless, a considerable underestimation of the SNP heritability is expected from trio data when oversampling multiplex families, or under assortative mating that has been confirmed for both ADHD and ASD (Chapter 6 Figure 1). Thus, I advise against the use of trio data for SNP heritability estimation, and advice to appropriately
scale the SNP heritability when applying data with unscreened controls for common disorders.

Assortative mating
Recently, Nordsletten et al published a comprehensive study indicating a clear pattern of nonrandom mating within and across eleven major psychiatric disorders based on over 700,000 psychiatric patients from the Swedish population. Although the partner-resemblances found were pronounced (e.g. a partner-correlation of 0.47 for autism), Nordsletten et al did not discuss the expected consequences for disorder prevalence and heritability in the offspring generation, despite available theory in the quantitative genetics literature. In Chapter 7, upper boundaries were explored of the consequences of assortative mating for psychiatric traits by applying quantitative genetic models. First of all, the correlation estimates of Nordsletten et al were found to likely overestimate the correlations in the full population, because they analyzed a study sample with oversampling of cases. For example, the partner-correlation in the full population for autism was approximated at around 0.28 rather than the 0.47 presented by Nordsletten et al. Based on several inevitable assumptions discussed in detail in Chapter 7, the heritability is likely to increase as a consequence of assortative mating with an upper boundary of around 5% for one generation of assortative mating and 13% for multiple generations. In addition, the population prevalence is expected to increase, with a more pronounced impact for strong assortative mating ($\rho$), for disorders with a low population prevalence ($K$), and high heritability ($h^2$). For example, a relative increase in population prevalence of around 50% would be expected for autism ($\rho = 0.28, K = 0.0015, h^2 = 0.8$) compared to only 0.6% for MDD ($\rho = 0.12, K = 0.15, h^2 = 0.35$) from one generation of assortative mating. I note that these numbers provide upper boundaries; when partner-resemblance would, for example, be partly attributable to social homogamy (partner-resemblance driven by shared environmental factors) rather than phenotypic assortment (partner-resemblance driven by the psychiatric trait), the consequences would be considerably less pronounced. Notably, assortative mating is not expected to affect GWAS results, because of the small proportion of variance explained by individual loci, as discussed in Chapter 6.

How can these modeled consequences be interpreted in the context of empirical data? First, it should be noted that the increase in prevalence is most pronounced for rare disorders, and a prevalence increase from parental to offspring generation of e.g. 1.0% to 1.5% would be hard to detect in empirical
data given the standard errors around these estimates. Second, the presented consequence provide upper boundaries; when the partner-resemblance found by Nordsletten et al would e.g. be partly attributable to social homogamy, the consequences would be considerable less. Third, other factors also affect the population prevalence, such as reduced fecundity in psychiatric patients, which might reduce the consequences of assortative mating to some extent. Notably however, assortative mating and reduced fecundity are not expected to balance each other in the long term: the consequences of assortative mating asymptotically reach equilibrium after several generations, whether the consequences of the reduced fecundity do not. As a general limitation, it should further be noted that modeling was based on several inevitable assumptions discussed in detail in Chapter 7. To conclude, the modeled consequences of assortative mating are difficult to test in empirical data, but suggest that the consequences of assortative mating are at most modest for the heritability but may be considerable for the population prevalence of rare psychiatric disorders. A challenge for future research will be to test what proportion of partner resemblance can be attributed to phenotypic assortment, secondary assortment, social homogamy and marital interaction. With the emergence of large-scale population based samples including genotyped spouse-data, opportunities may present to address this question.

Next to within disorder partner-resemblance, Nordsletten et al also found evidence for across-disorder partner-resemblance. Phillips et al explained that such across-disorder partner-resemblance may be attributable to across-disorder assortment, or to within-disorder assortment in addition to within-person correlation between both disorders. Nordsletten et al did not distinguish between these two scenarios, but Van Grootesteet al and found that across-trait partner-resemblance with respect to obsessive-compulsive, anxious and depressive symptoms was attributable to both across-trait assortment and within-person correlation. Notably, Wesseldijk et al found that across-trait partner-resemblance was more pronounced in parents of individuals affected with a psychiatric disorder. To model the genetic consequences of across-disorder partner-resemblance is complex, but it could be hypothesized that this phenomenon relates to the genetic correlation between different psychiatric disorders in one way or the other. First, the genetic correlation between psychiatric disorders can be hypothesized to reflect a general underlying liability for all (or several) psychiatric disorders, and individuals might get affected with e.g. either schizophrenia (SCZ) or bipolar disorder (BIP) based on environmental effects or disorder-specific genetic effects. When partners would assort based on
this general psychiatric liability, they might present as affected with different disorders assessed as across-disorder assortative mating. In other words, this first hypothesis could be thought of as “across-disorder partner-resemblance as consequence of assortment based on general psychiatric liability (detected as genetic correlation).” Nevertheless, another hypothesis might be in line with the opposite. Suppose that individuals would, indeed, mate based on distinct liabilities for actual different disorders, such as e.g. SCZ and BIP (now hypothesized entirely distinct disorders). Under this scenario, the risk alleles of those with high liability for SCZ would get together with the risk alleles for those with high liability for BIP, which would result in a correlation in the offspring generation between the effective loci on SCZ and the effective loci on BIP. This correlation between effective loci would be detected as genetic correlation between SCZ and BIP: in other words, the “genetic correlation would be attributable to across-disorder assortative mating.” I am not aware of methods suitable to distinguish between these two and other hypotheses, but this might be a challenge for future research.

Recent successes in psychiatric genetics
Since my PhD trajectory commenced in 2011, research on psychiatric genetics has made great progress. In 2015, the CONVERGE consortium identified and replicated two loci associated to MDD in a Chinese female sample comprising 5,337 screened controls and 5,303 cases with recurrent episodes of MDD. The success in this relatively small sample has been hypothesized to be attributable to the homogeneous cases, inclusion of female only and use of the same genotyping platform for the whole sample. This approach contrasts the more heterogeneous PGC sample including 9240 cases and 9519 controls that found no loci. Another recent success was obtained with a very different strategy: in a total of 180,281 individuals from European ancestry two loci were linked to depressive symptoms, which were heterogeneously assessed across contributing cohorts. The relevance of these latter findings for clinical diagnosis of MDD is not certain yet, but these findings may suggest that including enough individuals balances phenotypic heterogeneity. More interestingly, at the print of this thesis, the unpublished PGC-MDD meta-analysis comprising tens of thousands of cases from European ancestry also detected several of loci associated to MDD.38

The recent successes in genetic research on SCZ might hint to the successes to be expected in MDD, as research on SCZ seems to be a couple of years ahead of MDD (possibly owing to SCZ larger heritability, more homogeneous nature, and smaller population prevalence). In 2011, seven loci
had been detected for SCZ\textsuperscript{41} (which can roughly be compared to the current stage of genetic research in MDD), increasing to 108 loci identified in 2014 in a sample including 36,989 cases.\textsuperscript{42} Moreover, in 2016 Sekar et al found that the top finding for SCZ is likely to reflect different levels of complement component 4 (C4) in individuals with SCZ, while in mice C4 was found to mediate synaptic pruning in postnatal development.\textsuperscript{43} The finding of Sekar et al has been considered the first ‘inroad into the molecular etiology of SCZ’ that might potentially lead to new therapies in the future.\textsuperscript{44}

From a more skeptical point of view, one might wonder how much clinical relevance can be expected from genetic loci individually explaining typically less than 1\% of variance in disease risk. Professor Lander, first author of the Human Genome Project,\textsuperscript{45} replied to this in a masterclass that I attended in November 2015, and noted that statins are important drugs in preventing cardiovascular disease that also link to a SNP explaining very little variance; rs12916 in the HMG-CoA reductase gene explains only 1.6\% of variance in cholesterol levels.\textsuperscript{46} So, although SNPs that have been (and will be) identified for MDD explain only a very small proportion of variance in disease risk, they may still point to relevant biologically pathways that could possibly initiate novel pharmacotherapeutic development in the coming decades.

**Past challenges and the road to the future**

At the time of commencement of my PhD, research on genetic effects on MDD and other psychiatric disorders had been characterized by single study findings that were not replicated. Findings from linkage studies and candidate gene studies were inconsistent,\textsuperscript{47,48} and research on GxE with candidate genes has led to publication bias without any robust findings.\textsuperscript{2} Naturally, until recently large scale GWAS studies were not achievable, and in the candidate gene era it was like seeking a needle in a haystack without the appropriate instrument. This metaphor is further strengthened by the current knowledge about the small effect sizes of individual loci, that typically explain less than 1\% of variance in disease risk. The interaction between polygenic MDD risk and childhood trauma found in Chapter 4 can best be viewed as a single cohort phenomenon given the non-replication in Chapter 5. Personally, I do not think research on GxE analyses will be the best way forward for genetic research in MDD in the near future. Nevertheless, I would advise for GWAS cohorts to collect uniformly environmental information as much as possible to prepare for solid GxE testing in the further future.
Rather, for the next couple of years, I think progress can be expected from collaboration in the Psychiatric Genomics Consortium, and from the large-scale population samples, such as UK Biobank\(^49\) and iPSYCH. These population samples include many MDD cases based on MDD's high prevalence of around 15%, but are less suitable for rarer disorders such as schizophrenia with a prevalence of around 1%. The overall motto to me seems to regard single-study findings as hypothesis forming, and to aim for the largest possible sample size with subsequent independent replication. For the future, when sufficient causal loci have been identified, I think attempts can be made to integrate gene-findings with gene-expression and epigenetic information. This approach may help to improve classification of MDD and other psychiatric disorders from a more biologically informed point of view than the current DSM classification. I further anticipate that the rapidly increasing number of genetic correlation estimates will help to elucidate the relation between MDD, other disorders, and non-disease traits.

I think these are exciting times for genetic research in MDD. The GWAS sample size has increased drastically from 5763 cases by 2011\(^47\) to tens of thousands of cases in 2016, the exciting results of which are anticipated to be published shortly.\(^38\) In parallel, statistical methods were developed to effectively analyze this vast amount of data. An important method is LD score regression, which allows analyses of summary statistics to assess the variance explained by genotyped SNPs,\(^21\) and the genetic correlation between any two traits.\(^16\) It seems to me that the skepticism from a couple of years ago about the lack of GWAS findings,\(^50\) which had followed the initial excitement following completion of the Human Genome Project in 2001,\(^45\) has now been replaced again with more realistic optimism. I, for one, am looking forward to the years to come.

REFERENCES