

Nederlandse Samenvating

In **hoofdstuk 1** van dit proefschrift wordende methoden uiteengezet die ten grondslag liggen aan gedragsgenetisch onderzoek. In gedragsgenetisch onderzoek toetst met in hoeverre individuele verschillen in fenotypes worden bepaald door individuele verschillen in genotypes. Ieder complex menselijk gedrag, of aandoening, kan doorgaans worden onderzocht als fenotype. De oudste vorm van gedragsgenetisch onderzoek geschiedt door middel van tweelingstudies, die de mogelijkheid bieden om te onderzoeken in hoeverre individuele verschillen tussen mensen (de populatie variantie) worden verklaard door genetische factoren en door de omgeving. Tweelingstudies bieden de mogelijkheid om de erfelijkheid van een fenotype te onderzoeken omdat eeniige tweelingen hun volledige DNA delen, en twee-eiige tweelingen is dit bij benadering de helft van het DNA. Als eeniige tweelingen voor een bepaald fenotype meer op elkaar lijken dan twee-eiige tweelingen, dan is dit een aanwijzing dat individuele verschillen in dit fenotype door individuele verschillen in het genotype worden bepaald. Genetische factoren kunnen worden onderscheiden in “additief genetische effecten” en “non-additief genetische” of “dominantie effecten”. Bij effecten vanuit de omgeving kan een onderscheid worden gemaakt tussen omgevingsinvloeden die uniek zijn voor ieder kind binnen een tweelingpaar, en gedeelde omgevingsinvloeden, die in hun effecten hetzelfde zijn voor beide kinderen van een tweelingpaar. Op het moment dat tweelingonderzoek aantoont dat individuele verschillen deels verklaard worden door genetische factoren is vervolgens de vraag welke genen van invloed zijn.

Sinds de jaren '70 bestaat de mogelijkheid om door middel van genetische “linkage” (koppeling) de invloed van individuele genen of een individueel stuk genoom op gedrag te onderzoeken. Over het hele genoom delen twee-eiige tweelingen de helft van hun DNA, maar voor een bepaalde locus op het genoom zijn ze identiek, ze delen de helft, of ze zijn uniek. In koppelingsonderzoek wordt eerst bepaald of twee-eiige tweelingen, of broers en zussen, voor een bepaald stuk DNA identiek zijn, de helft van het DNA delen, of uniek zijn. Vervolgens wordt gekeken of de genetische overeenkomst op een bepaald stuk DNA tussen paren samengaat met overeenkomsten in gedrag, d.w.z. of de paren die het stuk genoom delen meer op elkaar lijken dan de paren die op dat stuk niets delen. Vanaf de jaren '90 wordt in de gedragsgenetica genoombreed koppelingsonderzoek uitgevoerd. Hierbij wordt op enkele honderden tot duizenden plekken over het hele genoom tegelijk geanalyseerd of verschillen in

het genoom samenhangen met verschillen in gedrag. Doordat men voor dit onderzoek broers en zussen gebruikt kan men met slechts enkele honderden tot duizenden metingen uitspraken doen over het hele genoom. Broers en zussen delen immers vaak lange aaneengesloten stukken genoom van een of beide ouders. Een eting van een enkel stukje DNA kan worden gegeneraliseerd over een langer stuk DNA, waar wel de zekerheid van de generalisatie afneemt naarmate de afstand met het gemeten stuk DNA toeneemt.

In 2000 werd de volledige sequentie van het menselijke genoom blootgelegd door het "Human Genome Project" wat de mogelijkheid bied om op een nog veel fijnere schaal genetisch onderzoek uit te voeren. In een genoombrede associatiestudie worden enkele honderdduizenden tot tientallen miljoenen "single nucleotide polymorphisms" (SNPs) onderzocht. De hoge dichtheid van de metingen betekent dat men deze associatiestudies ook kan uitvoeren in ongerelateerde individuen. Men is voor associatie niet meer afhankelijk van directe verwantschap tussen proefpersonen.

In de eerste genoombrede associatiestudies was het aantal significante associaties tussen gen en gedrag beperkt. Dit leidde tot de vraag of de veel voorkomende genetische varianten die in een dergelijke studie worden geanalyseerd wel van belang zijn bij dit soort fenotypes. Polygenetische analyses, waarbij het effect van al de genetische varianten gezamenlijk wordt onderzocht, liet zien dat deze SNPs wel degelijk een signaal herbergen, maar dat de effecten van de afzonderlijke SNPs zeer klein zijn. Dat betekent dat grote steekproeven nodig zijn om deze effecten te detecteren. Samenwerking tussen laboratoria over de hele wereld heeft inmiddels geleid tot de identificatie van honderden associaties tussen SNPs en menselijke fenotypes, zoals terug is te vinden in de GWAS catalogus (<https://www.genome.gov/26525384>). Het betreft bevindingen voor een groot aantal fenotypes zoals BMI, lichaamslengte, lipiden, hart- en vaatziekten, breinvolumes, maar ook psychiatrische ziekten zoals schizofrenie of bipolaire stoornis, en gedragsmaten zoals roken en neuroticisme.

Hoofdstuk 2 beschrijft een studie waarin met een longitudinaal tweelingmodel wordt gekeken hoe de genetische invloeden op symptomen van angst en depressie veranderen met leeftijd. Dit wordt ook wel "gen leeftijd interactie" genoemd. Hiervoor is een longitudinaal simplex model gebruikt, waarin de variatie op iedere leeftijd wordt gesplitst in genetische variatie en

omgevingsvariatie. De genetische en omgevingsvariaties worden verder opgedeeld in een component die op een eerdere leeftijd al invloed had op symptomen van angst en depressie en een invloed “innovatie” die nieuw is op deze leeftijd. Deze studie laat zien dat de genetische invloed op angst en depressie gedurende de levensloop toeneemt. Tot de leeftijd van 18 jaar is er dus steeds sprake van nieuwe genetisch effecten en deze toename stopt na het 18^e levensjaar. De omgevingsinvloeden op symptomen van angst en depressie nemen ook toe tijdens de jeugd. Deze toename van omgevingsinvloeden is (tijdens de jeugd) groter dan die van de genetische invloeden. Dit leidt ertoe dat de erfelijkheid, gedefinieerd als de genetische variatie gedeeld door totale variatie, afneemt van de vroege jeugd tot de volwassenheid. Er is geen evidentie dat nieuwe genetische effecten op volwassen leeftijd invloed uitoefenen op symptomen van angst en depressie, terwijl er wel evidentie is voor nieuwe omgevingsinvloeden op volwassen leeftijd.

Hoofdstuk 3 beschrijft een studie waarin op populatieniveau groeicurven worden geschat voor de op DSM-IV gebaseerde internaliserende en externaliserende stoornissen. Deze groeicurven beschrijven de in deze populatie meest voorkomende ontwikkelingstrajecten voor externaliserend gedrag en internaliserend gedrag tussen leeftijd 7 en 18. Uiteindelijk beschrijft het model waarin er 5 discrete ontwikkelingstrajecten bestaan voor zowel externaliserend als internaliserend gedrag deze data het beste. Voor de internaliserende stoornissen is er evidentie voor trajecten met een “blijvend zeer lage symptoom score”, een “lage symptoom score”, “stijgende symptoom score”, “afnemende symptoom score” en een “in de adolescentie toenemende symptoom score”. Voor externaliserende stoornissen zijn er ook de eerste vier trajecten. In plaats van een traject van de “in de adolescentie toenemende scores” is er evidentie voor een traject met “blijvend hoge externaliserende scores”. Covariaten gemeten rond de zwangerschap, zoals bijvoorbeeld het roken van de moeder, verhogen de kans om in een hoge of stijgende externaliserende klasse terecht te komen, ook mannen zitten vaker in een stijgende of hoge externaliserende klasse. Vrouwen hebben een verhoogde kans om in een hoog scorende internaliserend traject te worden geclassificeerd. Na de internaliserende en externaliserende trajecten te hebben vast gesteld, onderzoeken we comorbiditeit tussen internaliserende en externaliserende trajecten. Hier kijken we of bepaalde internaliserende trajecten vaker dan verwacht samenhangen met bepaalde externaliserende trajecten.

Vergelijkbare trajecten komen vaker dan verwacht samen voor: kinderen die in toenemende mate internaliserende problemen hebben gaat vaak samen met toenemende externaliserende problemen, hetzelfde geldt voor afnemende problemen. Het internaliserende traject voor individuen die in de adolescentie een stijgende score laten zien blijkt onafhankelijk van de hoge en stijgende externaliserende trajecten. Deze studie laat zien dat eventuele comorbiditeit moet worden overwogen als een kind of adolescent wordt gediagnosticeerd met een externaliserende of internaliserende stoornis, niet alleen ten tijde van het stellen van de diagnose, maar ook later in het traject.

Hoofdstuk 4 beschrijft een genomebrede associatie (genome wide association: GWA) studie naar internaliserende problemen op 3 jarige leeftijd, uitgevoerd in 4596 kinderen. De meta-analyse is gebaseerd op analyses in drie cohorten: het Nederlandse Tweeling Register, “generation R” (een Rotterdamse populatiestudie) en “The Western Australian Pregnancy Cohort” (RAINE, een Australische populatiestudie). Zoals op grond van het aantal proefpersonen verwacht kan worden laat de meta-analyse geen genomebrede significante associaties zien. Wel zijn er suggestieve signalen in het PCSK2 gen dat zich bevindt op chromosoom 20p12.1. Overige analyses laten zien dat alle SNPs gezamenlijk tussen de 13% en 43% van de individuele verschillen in internaliserend gedrag van 3 jarigen verklaren. Verder blijken SNPs die in eerdere studies zijn geassocieerd met psychiatrische fenotypes in kinderen en volwassenen ook een collectief effect te hebben op internaliserend gedrag in 3 jarigen. Hoewel deze studie geen individuele varianten identificeert, is wel evidentie vergaard dat de genetische invloeden op internaliserende problemen bij 3 jarigen zeer poly-genetisch zijn, dat wil zeggen dat er veel genen zijn met ieder een zeer klein effect. Dit is in lijn met bevindingen uit grotere meta-analyses van psychiatrische fenotypes bij volwassenen.

In **Hoofdstuk 5** is gekeken naar de overlaptussen het genetisch risico op schizofrenie en “DSM-IV gebaseerde probleemscores” bij Nederlandse en Britse kinderen. Op basis van een genomebrede meta-analyse naar schizofrenie bij meer dan 30.000 schizofreniepatiënten en meer dan 40.000 gezonde controle-individuen werd een lineaire combinatie van SNPs gecreëerd. Deze combinatie van SNP's voorspelt schizofrenie in onafhankelijke cohorten (genetisch profiel score: polygenic risk score: PRS). Dezelfde PRS wordt gebruikt om symptomen te

voorspellen voor DSM-IV depressie, angststoornissen, externaliserende stoornissen en ADHD op leeftijd 7, 10, 13 en 15 in Britse en Nederlandse kinderen. Meta-analyse van de resultaten laat zien dat de PRS voor schizofrenie samenhangt met angst op leeftijd 10 en in mindere mate met angst op leeftijd 7 en 13 en depressie op leeftijd 7 en 10. Post-hoc toetsen laten zien dat er een sterkere samenhang is tussen PRS en internaliserende problemen in kinderen dan tussen PRS en externaliserende problemen in kinderen.

In Hoofdstuk 6 wordt een associatie tussen het catenin delta 2 gen (CTNND2) en angst die gevonden is in ratten, gerepliceerd in mensen. Data van het internationale "Psychiatric Genetics Consortium"(PGC) en de "Nederlandse studie van angst en depressie" (NESDA) zijn hiervoor geanalyseerd. Hieruit blijkt dat de varianten in het CTNND2 gen, collectief, meer met schizofrenie en depressie zijn geassocieerd dan verwacht op basis van kans. Voor angst en bipolaire stoornissen werd geen associatie aangetoond met CTNND2. Deze studie laat zien dat het waardevol kan zijn om bevindingen uit proefdierstudies te repliceren.

In **Hoofdstuk 7** wordt een nieuwe statistische techniek beschreven om gen-omgevingsinteracties te onderzoeken en te kwantificeren. Deze techniek maakt gebruik van een genetische relatie matrix (GRM). Deze matrix bevat voor een grote groep personen hun genetische verwantschap met alle andere personen in de studie voor wie SNPs gemeten zijn. Voor ongerelateerde personen is de verwachte verwantschap geschaald naar 0, voor broers en zussen is de verwachte verwantschap 0.5, maar rondom deze verwachtingen is een zekere mate van spreiding. Sommige ongerelateerde personen delen meer SNPs dan andere, en de vraag is of ze ook fenotypisch meer op elkaar lijken. De verwantschappen in de GRM worden bepaald op basis van gemeten SNPs. De GRM matrix wordt vervolgens gesplitst. Er is 1 matrix die alle informatie bevat en een tweede die slechts de verwantschap informatie bevat voor mensen met een paarsgewijze verwantschap boven de 0.05, dus van mensen die aan elkaar gerelateerd zijn. Door vervolgens de variantie in het fenotype te modelleren als een functie van beide matrices, kan men bepalen welk deel van de variantie in het fenotype verklaard wordt door de gemeten SNPs, en welk deel door alle genetische effecten. Ik heb vervolgens een methode ontwikkeld, die moderatie van de genetische variantie toestaat door een continue verdeelde variabele. Uit de analyses van BMI, lichaamslengte, symptomen van angst en depressie en attentieproblemen, blijkt dat de erfelijkheid van BMI en attentieproblemen

varieert met leeftijd, die van lichaamslengte met geboorte jaar. Bij BMI komt dit door veranderingen in de genetische invloeden en de omgevingsinvloeden op het fenotype, bij lichaamslengte en attentie problemen komt dit door een verandering in de omgevingsvariantie. Er is geen moderatie van de genetische effecten van angst en depressie.

Hoofdstuk 8 van dit proefschrift beschrijft een methode om de genetische covariantie tussen meerdere fenotypes vast te stellen op basis van gemeten SNPs en verwantschapsinformatie. Hier wordt zoals in hoofdstuk 7 een GRM gebruikt. Voor het schatten van de genetische covariantie tussen meerder fenotypen wordt de som van twee fenotypen gebruikt. De variantie van de som van 2 variabelen is immers de variantie van de eerste variabele plus de variantie van de tweede variabele plus twee maal de covariantie tussen de twee variabelen. Door de variantie van de som van twee variabelen in genetische en omgevingscomponenten te ontbinden en de twee afzonderlijke variabelen in een genetische en omgevingscomponent te ontbinden kan men de genetische en omgevingscovariantie tussen de twee variabelen bepalen. Met deze methode is een computationeel intensieve stap in multivariaat modeleren te voorkomen. De hier voorgestelde techniek is toegepast om de genetische covariantie te schatten tussen de 12 neuroticisme items en de 12 extraversie items van de NEO persoonlijkheidsvragenlijst. Deze analyses laten zien dat de covariantie structuur tussen deze items, gebaseerd op gemeten genetische variatie, een duidelijk onderscheid maakt tussen de neuroticisme items en de extraversie items. Hiermee is de op basis van tweelingstudies voorspelde covariantie structuur teruggevonden in moleculaire genetische data.

Hoofdstuk 9 en **Hoofdstuk 10** betreffen de Engelstalige samenvatting, en discussie van het proefschrift. In de discussie van dit proefschrift wordt een breder kader van de humane genetica geschetst.