

# **Amyotrophic Lateral Sclerosis: risk factors in the genome and exposome**

**Perry van Doormaal**

**Promotie 21 november 2017, UMC Utrecht**

**Titel proefschrift: Amyotrophic Lateral Sclerosis: risk factors in the genome and exposome**

## **Samenvatting in het Nederlands**

In dit proefschrift wordt het onderzoek beschreven naar genetische factoren en omgevingsfactoren die effect hebben op het risico voor het ontwikkelen van amyotrofische laterale sclerose (ALS).

ALS is een progressieve neurodegeneratieve ziekte die met name de motorische zenuwen (motorneuronen) aantast in zowel de hersenen, ruggenmerg als de perifere zenuwen. Klinisch wordt de ziekte dan ook gekenmerkt door een progressieve zwakte en dunner worden van de spieren waarbij ook spasticiteit ontstaat. ALS is een ongeneeslijke ziekte waarbij het enige bewezen medicijn, riluzole, slechts een beperkte remming op de ziekteprogressie laat zien. De meeste patiënten overlijden aan respiratoire zwakte, maar de tijd van de eerste symptomen tot overlijden is hoogst variabel. De helft van de patiënten is na 34 maanden overleden maar 20 procent overleeft vijf tot tien jaar na het begin van de symptomen. In **hoofdstuk 1** wordt hiernaast besproken dat in de afgelopen jaren is ontdekt dat ook gebieden buiten de motorneuronen aangedaan zijn.

In de minderheid van de patiënten (10 tot 15%) komt de ziekte in de familie voor. Hierbij wordt de ziekte veroorzaakt door een sterke genetische oorzaak met hoge penetrantie. Voor de resterende groep patiënten wordt aangenomen dat de ziekte veroorzaakt wordt door een combinatie van factoren, zowel genetisch als in de omgeving van de patiënt, die de kans op de ziekte vergroten. Tweelingstudies laten zien dat de genetische bijdrage hierbij het grootst is, 61%. Maar omgevingsfactoren leveren dus, met een risicobijdrage van 39%, ook een substantiële bijdrage.

Deze risicofactoren leren kennen is van zeer groot belang. Dit is niet enkel omdat we zo meer te weten komen over de ziekte, maar ook omdat deze factoren aangrijpingspunten kunnen zijn voor gerichte

## **Amyotrophic Lateral Sclerosis: risk factors in the genome and exposome**

behandelingen. Daarom is het doel van mijn proefschrift om risicofactoren te ontdekken op het gebied van genetica en omgeving, die invloed hebben op ALS.

### **Omgevingsfactoren**

In de laatste jaren zijn er meerdere aandoeningen gevonden die geassocieerd zijn met ALS. Zo blijkt dat frontotemporale dementie (FTD) frequent voorkomt samen met ALS en er een duidelijke genetische overlap is tussen deze twee aandoeningen met gemeenschappelijke mutaties in onder andere *C9orf72*, *TDP-43* en *VCP*. Mogelijk zijn er meer aandoeningen die met ALS geassocieerd zijn. In een Brits tertiair verwijzingscentrum voor ALS werden eerder zeven patiënten beschreven die voorafgaand aan de diagnose ALS reeds gediagnosticeerd waren met multiple sclerosis (MS). In vier van de vijf patiënten waarvan DNA beschikbaar was, werd een repeat-expansie gevonden in het *C9orf72* gen. Er werd gesteld dat het gezamenlijk voorkomen van ALS en MS vaker voorkomt dan statistisch verwacht, en dat een *C9orf72* repeat-expansie hieraan ten grondslag ligt. Echter door de kleine aantallen patiënten en doordat deze patiënten allen verwezen zijn naar een tertiaire kliniek, met mogelijk referral bias tot gevolg, is deze associatie zonder replicatie onvoldoende bewezen. In **hoofdstuk 2** onderzoek ik de associatie tussen ALS en MS in een population-based case-control studie in Nederland, waar middels vragenlijstonderzoek 810 ALS patiënten hun medische voorgeschiedenis beschrijven. Drie patiënten beschreven zelf dat ze ooit gediagnosticeerd zijn met MS, waarbij de medische status dit in twee patiënten bevestigde (0.25% van de patiënten). Dit vóórkomen van MS in deze ALS populatie is niet anders dan wat verwacht kan worden op basis van het vóórkomen in de Nederlandse bevolking. Ook hebben deze twee patiënten geen verlengde repeat-expansie in *C9orf72*. Ik concludeer hier dan ook dat, in tegenstelling tot eerdere bevindingen, de diagnose MS niet vaker dan verwacht gediagnosticeerd is bij Nederlandse ALS patiënten.

In **hoofdstuk 3** onderzoek ik of andere diagnoses wel geassocieerd zijn met ALS en of medicatie hierbij ook een relatie tot de ziekte kent. Middels vragenlijsten beschrijven 722 patiënten en 2268

## Amyotrophic Lateral Sclerosis: risk factors in the genome and exposome

controlepersonen, gematcht op leeftijd en geslacht, hun medische voorgeschiedenis en medicatiegebruik.

Hoofdtrauma in het verleden blijkt vaker voor te komen bij patiënten dan bij controles en lijkt een risicofactor (odds ratio (OR) 1,95;  $P = 0,02$ ). Daarentegen komt hypercholesterolemie (OR = 0,76;  $P = 0,006$ ), statine gebruik (OR = 0,45;  $P = 1,86 \times 10^{-9}$ ) en immunosuppressiva gebruik (OR = 0,26;  $P = 0,03$ ) minder vaak voor bij patiënten en dit lijken dus beschermende factoren.

Een nadeel van dergelijk hypothese-gestuurd onderzoek is dat er slechts naar één categorie omgevingsfactoren wordt gekeken en geen rekening wordt gehouden met het geassocieerde effect van andere omgevingsfactoren. Zo zou roken een effect kunnen hebben op zowel het risico van ALS, als op het voorkomen van hypercholesterolemie. Hierdoor zou hypercholesterolemie een correlatie met ALS kunnen tonen, terwijl de echte oorzakelijke factor roken kan zijn. In **hoofdstuk 4** beschrijf ik een nieuwe methode om naar omgevingsfactoren te kijken. Dit betreft een exposoom-wijde associatie studie, waarmee voor grote groepen omgevingsfactoren gekeken kan worden naar het onafhankelijke effect van (groepen geassocieerde) omgevingsfactoren. Hiervoor maak ik gebruik van informatie over 133 omgevingsfactoren van 1330 patiënten en 3347 gematchte controles die middels LASSO analyse en principale componenten regressie worden geanalyseerd. Beschermende principale componenten zijn: 'voedingspatroon met veel antioxidanten en lage energy inname' (OR = 0,73;  $P = 7,24 \times 10^{-17}$ ); 'voedingspatroon met veel polysaccharide, plantaardige eiwitten en koper, en weinig verzadigde vetzuren' (OR = 0,89;  $P = 2,53 \times 10^{-4}$ ); 'hoog cardiovasculair risicoprofiel' (OR = 0,93;  $P = 0,048$ ); en 'alcoholgebruik ten tijde van het begin van de ziekte met voedingspatroon met weinig retinol' (OR = 0,68;  $P = 2,54 \times 10^{-30}$ ). Risicofactoren zijn: 'kleine lengte en laag gewicht' (OR = 1,54;  $P = 1,37 \times 10^{-21}$ ); 'gewichtveranderingen' (OR = 1,11;  $P = 0,0015$ ); 'roken ten tijde van het begin van de ziekte met een lage ontbijtfrequentie' (OR = 1,18;  $P = 4,06 \times 10^{-7}$ ); en 'voedingspatroon met veel glutamaat en hoog energetische hobby's' (OR = 1,12;  $P = 3,82 \times 10^{-4}$ ). Deze methodiek heeft als voordeel dat de onafhankelijke effecten op het ontstaan van een ziekte worden bepaald en er rekening wordt gehouden met de vele factoren die er getest worden. Deze methode kan ook voor andere complexe aandoeningen

## **Amyotrophic Lateral Sclerosis: risk factors in the genome and exposome**

gebruikt en aangepast worden. Op deze manier is een stap gezet om van hypothese-gestuurd onderzoek naar één specifieke groep omgevingsfactoren de omslag te maken naar hypothese-vrij onderzoek naar het exposoom, het totaal aan omgevingsfactoren dat een persoon tegenkomt in het leven.

### **Genetische factoren**

In de genetica is deze stap naar hypothese-vrij onderzoek reeds gemaakt. Waar eerder gen-gedreven studies de norm waren, is in het afgelopen decennium de nadruk komen te liggen op genoom-wijde associatie studies (GWAS) en whole-genome of whole-exome sequencing. De resultaten van een GWAS bleken echter niet altijd goed replicateerbaar en het aantal fout-positieve resultaten lag hoog, met name in de beginjaren van GWAS toen de methodiek nog niet gestandaardiseerd was. Voor ALS zijn meerdere GWAS uitgevoerd, waarbij in een van deze studies er een overlevingsvoordeel werd gevonden voor een single nucleotide polymorphism (SNP) in het *KIFAP3* gen. In **hoofdstuk 5** beschrijf ik een replicatiestudie waarbij deze SNP is bepaald in 2362 nieuwe patiënten, waarbij 8,1 procent van hen het geassocieerde genotype had. Dit genotype liet in geen van de verschillende overlevingsanalyses een significant effect op overleving zien. Hoewel een populatie-specifiek effect niet uit is te sluiten, lijkt het positieve effect op overleving van deze SNP, ook gezien drie inmiddels gepubliceerde negatieve analyses in andere populaties, niet aannemelijk.

**Hoofdstuk 6** beschrijft de uitkomsten van het next-generation sequencing onderzoek naar *de novo* mutaties in patiënten met ALS. Whole-genome en whole-exome data van 82 patiënten en hun beide gezonde ouders (in totaal 246 genomen) uit vier Europese landen werd geanalyseerd op mutaties die bij de patiënt maar niet bij hun ouders gevonden werden (zogenaamde *de novo* mutaties). Hierbij werden 69 *de novo* mutaties gevonden, resulterend in gemiddeld 0,84 *de novo* mutaties per patiënt. Dit is niet hoger dan in de algemene populatie. Ook na combinatie met alle hiervoor gepubliceerde *de novo* mutaties uit eerdere studies was dit gemiddelde (0,64) niet afwijkend van de algemene populatie. Er waren geen overlappende *de novo* mutaties en de genen waarin deze mutaties gevonden werden, waren niet betrokken

## **Amyotrophic Lateral Sclerosis: risk factors in the genome and exposome**

bij eerder geopperde eiwitmechanismen. Ook was er geen onderlinge eiwit- interactie of interactie met bekende ALS eiwitten. De genen waarin de *de novo* mutaties bij ALS patiënten voorkwamen, bleken genen te zijn met een hoge kans op *de novo* mutaties in de algehele bevolking. Hoewel een effect van een enkele *de novo* mutatie op een specifiek plaats in het genoom niet uitgesloten is, hebben *de novo* mutaties geen grote rol in de ontwikkeling van ALS.

In **hoofdstuk 7** wordt de ontdekking van *NEK1* als ALS risicogen beschreven. Er werden 1022 index familiale ALS patiënten en 7315 controles whole-exome gesequenced. Een nieuwe screeningsstrategie, waarbij een burdenanalyse werden getraind met reeds bekende ALS genen, liet zien dat loss-of-function varianten in *NEK1* significant geassocieerd zijn met ALS. Dit resultaat kon worden gerepliceerd in 3362 sporadische ALS patiënten (OR = 8,8;  $P = 3.4 \times 10^{-8}$ ). Hiernaast vonden we in een genetisch gesloten gemeenschap in Nederland middels homozygosity-mapping dat de *NEK1* p.Arg261His variant een risicofactor leek te zijn voor ALS. In 7194 ALS patiënten en 4417 controlepersonen uit acht landen werd dit effect van deze mutatie bevestigd (OR = 2,4;  $P = 1.2 \times 10^{-7}$ ). Aangezien risicovarianten in drie procent van de patiënten optreden, is *NEK1* een relatief vaak voorkomende risicofactor voor het ontstaan van ALS.

In een Amerikaans cohort van ALS patiënten waarbij ALS in de familie voorkomt, zijn mutaties in *UBQLN2* gevonden als oorzaak voor ALS met hoge penetrantie in families. Uit verschillende replicatiestudies in verschillende populaties bleek dit een weinig voorkomende oorzaak te zijn voor ALS. In **hoofdstuk 8** beschrijf ik de analyse van het *UBQLN2* gen in Nederlandse families met ALS. Er werden geen pathogene mutaties gevonden in deze groep. Mutaties in *UBQLN2* zijn niet gevonden als oorzaak voor ALS in Nederland.

Bij patiënten met familieleden met ALS werd eerder gerapporteerd dat er vaker mutaties in meerdere ALS-geassocieerde genen gevonden worden in één persoon. Dit aantal is groter dan enkel door toeval te verklaren is. **Hoofdstuk 9** beschrijft hoe dit ook geanalyseerd wordt in sporadische ALS patiënten. In 4,1

## **Amyotrophic Lateral Sclerosis: risk factors in the genome and exposome**

procent van 755 patiënten en in 1,3 procent van 959 controles worden dergelijke dubbelmutaties gevonden, wat niet significant verschillend is. Wel zien we significant vaker een mutatie in ALS geassocieerd genen (*ATXN2*, *NIPAI*, *SMN1*) bij dragers van een repeat-expansie in het *C9orf72* gen ( $P = 0.001$ ), wat voornamelijk te verklaren is door aanwezigheid van een *NIPAI* repeat naast de repeat-expansie in *C9orf72* ( $P = 0.006$ ). Dit resultaat kan duiden op een rol van *NIPAI* in de variatie van ziektemanifestatie (zoals beginleeftijd of overleving) bij patiënten met een *C9orf72* repeat-expansie.

In **hoofdstuk 10** bespreek ik de hierboven beschreven studies, plaats ik deze in een breder kader en geef mogelijkheden voor toekomstig onderzoek.