Study on the relationship between genes and disease severity, disease burden and effectiveness of disease modifying medication in Dutch patients with Multiple Sclerosis

Published: 14-03-2017
Last updated: 12-04-2024

We aim to use the results from genome-wide association studies (GWASs) of phenotypic, disease-related data of individual MS patients to provide insights into genes contributing to disease severity and burden (primary objective), and effectiveness of...

Ethical review	Approved WMO
Status	Recruitment stopped
Health condition type	Autoimmune disorders
Study type	Observational invasive

Summary

ID

NL-OMON45439

Source
ToetsingOnline

Brief title
GEMS study - Genetic Etiology of Multiple Sclerosis

Condition

- Autoimmune disorders
- Demyelinating disorders

Synonym
demyelinating disease, multiple sclerosis

Research involving
Human

Sponsors and support

Primary sponsor: Drug Target ID, Ltd.
Source(s) of monetary or material Support: subsidie door Drug Target ID;Ltd. (www.drugtargetid.nl)

Intervention

Keyword: disease burden, disease severity, genetic factors, multiple sclerosis

Outcome measures

Primary outcome

Here, we first define the genotypic and phenotypic measurements. Subsequently, we will provide the parameter settings to perform the gene set analysis and PRS analysis that we will perform on the measured data.

Genotypic/GWAS data

For each patient we will establish, from the blood sample, a so-called genotype (one out of the two possible SNP variants) for the 550.000 SNPs on the GWAS chip. We will then use these genome-wide genotyping data and the collected phenotypic data to conduct a *within case* GWAS of disease severity, disease burden and the effectiveness of a number of selected DMTs for MS (see above). The resulting GWAS data will consist of a P value that indicates the likelihood of association between each of the 550.000 SNPs and each of the phenotypes

Phenotypic data

Using the outcomes of the online questionnaires (MSQoL-54, MSIP and General Assessment) and the neurological data (EDSS) we will derive the following phenotypic values for each patient:

Neurologist-derived disease severity. The MS Severity Score (MSSS) can be obtained by referencing the EDSS and diseases duration (the time between disease onset and time of EDSS scoring) to the global MSSS Table (Roxburgh et al., 2005). The MSSS Table essentially ranks individuals from lowest EDSS to highest EDSS for a given disease duration, and expresses this as a decile rank between 0 (least affected) and 10 (most severely affected). The MSSS will be computed using the MSSS test software version 2.0 (Roxburgh et al., 2005).

Patient-derived disease severity. The MSIP is a questionnaire to measure patient-reported severity of MS-related disability (Wynia et al., 2008). The MSIP consists of 36 items divided over seven scales and has four additional impairment items (symptoms). Item scores are graded on three to five-point rating scales with discrete responses, ranging from 0 (no disability) to 3 or 4 (complete disability). To assess patient-derived disease severity the respective scores are related to disease duration. Whereas the EDSS score reflects overall disability, the MSIP scales enable the calculation of disease severity in terms of functional systems and symptoms, which might be important given the heterogeneity of symptoms and severity in MS.

Patient-derived disease burden. The MSQoL-54 is a questionnaire to measure the health-related quality of life in patients with MS. It is composed of 54 items with item scales ranging from 1 to 5 , with a lower score indicating a higher impact on the quality of life, and vice versa (Acquadro et al., 2003); the MSQoL-54 provides a physical and a mental domain score.

In addition to measuring the objectified impact of MS, the MSIP also measures the subjective burden the patient experiences from the respective disabilities

(Wynia et al., 2008). Thus the MSIP also enables the assessment of the patient-derived disease burden by relating the scores to disease duration.

Disease Modifying Treatment (DMT). DMT assessment will include nine categories: interferon-beta (five products), glatiramer acetaat (two products), natalizumab, fingolimod, dimethyl fumarate, teriflunomide, alemtuzumab, daclizumab and *no medication*. We will determine the current and past DMT use through the General Assessment questionnaire. For each DMT used, the patient will be asked to rank, on a scale from 0 to 10, how they perceived its the effectiveness of the DMT to (1) reduce the number of relapses and (2) reduce disease symptoms. Based on these medication data, we will decide for which DMTs we have enough case subjects to conduct within-case GWASs of their effectiveness.

Missing data

The MS4 Research Institute will collect phenotypic data via online questionnaires; a patient can only participate in the blood sample collection for the genetic part of the study when he or she has completed the online questionnaires. Missing clinical data will be collected together with the neurologist as completely as possible; in the case of missing EDSS scores, a trained nurse will contact the patient to partake in a Telephone EDSS scoring.

Primary study parameter(s)

In order to perform gene set analysis and polygenic risk scoring, we first perform several preprocessing steps using standardized software for genetic

analyses. We will impute any missing genotypes using IMPUTE2 version 2 software (Howie et al., 2009) and we will perform data clumping to account for linkage disequilibrium (LD) blocks in the genome using PLINK version 1.9 or higher (Purcell et al., 2007). Subsequently, to perform GWAS on quantitative phenotypes we will use SNPtest version 2.5.2 or higher (Marchini and Howie, 2010). For the primary objectives, the three quantitative phenotypes are: clinically-derived disease severity, patient-derived disease severity and disease burden. These GWAS results can then be used to perform gene set analysis and polygenic risk scoring for which we describe the parameter settings in detail below.

Gene set analysis

We will test whether the genes in a gene set were jointly associated with the phenotype, for methods see (Bralten et al., 2013; Naaijen et al., 2017; Poelmans et al., 2011b). In brief, we will use a competitive test that examines whether a certain gene set of interest is more strongly associated with a phenotype than all other genes in the genome, correcting for gene size and density (Naaijen et al., 2017). The effect of the gene set is compared with the background signal of all genes that are not in the respective gene set. We will apply Bonferroni correction for the number of phenotypes tested.

Polygenic risk scoring

To test our hypothesis derived from the molecular landscape that a genetic predisposition for low levels of vitamin D is genetically related to MS, we

calculate the PRS between (A) genetic variants related to vitamin D regulation, for which the genetic data is publicly available (Wang et al., 2010) and (B) the genetic variants related to the GWASs for the three quantitative phenotypes, for which the genetic data is collected in this study. To obtain PRS values, we use the PRS analysis tool PRSice (Euesden et al., 2015). PRS values are the sum of SNPs associated with MS weighted by their effect sizes estimated from the vitamin D SNPs, from which only the SNPs exceeding seven broad P-value thresholds (0.001, 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5) are used. As such, seven PRS are generated based on all SNPs. In a similar manner, we will test the shared genetic etiology for another trait, the regulation of lipids/lipoproteins for which the data is also publicly available (Willer et al. 2013 , Ruth et al., 2016).

PRS analyses includes multiple comparisons (here we obtain 6 PRS outcomes which each contain 7 thresholds), which has to be corrected for. The p-values are calculated from a shared underlying (genetic) data set and therefore the use of Bonferroni correction would be overly conservative. To correct for multiple testing, we will aggregate the p-values and use the false discovery rate (FDR) method, incorporating potential dependencies between p-values (Glickman et al., 2014). Recent publications have used the same approach, for example Harris et al. aggregated p-values and used FDR correction on PRS analysis for 21 traits (with 7 thresholds) (Harris et al., 2016) and Hagenaars et al. (Hagenaars et al., 2016) used multiple cognitive test outcomes and 470 PRS p-values values were aggregated and FDR-corrected. We chose to aggregated all p-values from the six traits and apply FDR-correction on all 42 p-values simultaneously; the

advantage of aggregating the p-values is that the q-values, which are essential to perform FDR-correction, can be better estimated due to a larger amount of samples in the p-value distribution.

Secondary outcome

To answer our primary research question, we perform gene set analysis and PRS analysis on the phenotypic data which are continous variables. Here, to answer our secondary research question, we will perform the same analysis routines on the effectiveness of the selected DMTs (categorical data). These analysis have a more explorative nature, because it is currently difficult to estimate the medication use in this patient cohort.

Study description

Background summary

1.1 Multiple sclerosis
MS is a chronic disease affecting the CNS and is caused by a complex interplay of genetic and environmental factors. In MS, inflammatory and neurodegenerative processes cause damage to nerve cells (or neurons), which eventually leads to a loss of the electrically isolating myelin sheet around the axons (i.e. the main extensions of neurons that also communicate with other neurons through synapses). In turn, this loss of myelin and hence neuronal functionality gives rise to a variety of neurological symptoms and impairments, depending on the brain region where the demyelination occurs.
Most patients with MS initially experience periods of weeks to months with more or less extensive neurological symptoms and impairments (relapses), followed by periods in which these symptoms/impairments disappear partially or completely; this is called the relapsing-remitting form of MS (RRMS). In the absence of DMT use, more than 50% of patients with RRMS will develop progressive disability after approximately 15 years (secondary progressive MS, SPMS) (Rovira et al., 2013). Therefore, appropriate treatment planning during the early stage of the disease is of critical importance to optimize treatment outcomes and hence the overall prognosis of the disease.

Currently, thirteen DMTs are approved and registered for the treatment of RRMS,

namely interferon-beta (five products), glatiramer acetaat (two products), natalizumab, fingolimod, dimethyl fumarate, teriflunomide, alemtuzumab and daclizumab. Effective MS treatments reduce the number and severity of relapses, and of disease burden. Despite the broadening range of available treatments, the response of MS patients to - and therefore the effectiveness of - DMTs remains unpredictable and heterogeneous (Freedman and Abdoli, 2015; Serana et al., 2014). Furthermore, it has been suggested that genetic heterogeneity influences the pathogenesis of disease, and is involved in the disease progression, i.e. the number of relapses, the rate of disease progression and the overall disease burden (Baranzini et al., 2002; Freedman and Abdoli, 2015).

Pharmacogenetics and personalized treatments
In general, recent developments in the field of so-called *omics* data (genomics, transcriptomics, proteomics and metabolomics), have shown to be a promising strategy to predict the response to a given treatment. The data gathered through 'omics' approaches can therefore contribute to personalize the treatment plan of individual patients (Bhargava and Calabresi, 2016; Comabella and Vandenbroeck, 2011; Farias and Santos, 2015).
Pharmacogenetic studies aim to identify genetic variations - in the form of single nucleotide polymorphisms (SNPs) - that influence the efficacy of drugs, and thus the effectiveness of drug treatment. IFN-* for example, one of the most common first line treatments of MS, has no effect on 30-50% of MS patients (Kulakova et al., 2012). In addition, specific SNPs were shown to have a predictive value on the effectiveness of IFN-* treatment (Hoffmann et al., 2008; Kulakova et al., 2012; Wergeland et al., 2005).
For pharmacogenetic studies, it is critically important to couple genetic variations to well defined phenotypes (Mahurkar et al., 2014). In this respect, several questionnaires exist that capture the important clinical parameters and variation in disease presentation such as the Multiple Sclerosis Severity Score (MSSS) and the Multiple Sclerosis Impact Profile (MSIP), which assess disease severity, and Multiple Sclerosis Quality of Life-54 (MSQOL-54) questionnaire, which assesses disease burden (Roxburgh et al., 2005; Zettl et al., 2014).

Big *omics* data for complex diseases are generally analysed using bioinformatics-based tools and computational modelling, which leads to so-called *molecular networks*. We developed an innovative and unique method to map the molecular processes involved in complex diseases, i.e. building a so-called *molecular landscape*.
Building a molecular landscape entails the use of bioinformatics-based tools to analyse data from genome-wide association studies (GWASs) and genome-wide expression data (i.e. programs such as Ingenuity [Qiagen, Redwood City,US] for searching/identifying protein-protein interactions and gene enrichment) together with an extensive systematic review of biomedical literature on the (cell) physiology and pathological processes involved in complex diseases.
Through our method, we have built molecular landscapes for several neurological diseases such as Parkinson*s disease, Amyotrophic Lateral Sclerosis (ALS), Alzheimer*s disease and also for psychiatric disorders such as Attention

Deficit Hyperactivity Disorder (ADHD), autism, and obsessive compulsive disorder, which have already been published in high-ranked medical journals (Klemann et al., 2016; Poelmans et al., 2011a, 2011b, 2013; Vallès et al., 2014; van de Vondervoort et al., 2016).
Using our approach, we built a molecular landscape of MS, based on genetic data from nine publicly available GWAS data sets, which in total contain data for 11,500 MS patients and 21,000 healthy controls. Furthermore, we included genome-wide expression data from MS patients and data from MS cell and animal models. The molecular MS landscape is mainly located in neuronal cells and regulates a number of distinct biological processes and signalling cascades that contribute to key neuronal functions such as neuronal growth and (re)myelination. A number of signalling cascades within the landscape are directly linked to some of the existing etiological hypotheses of MS, such as disturbances in vitamin D signalling, neuronal proliferation and differentiation, and signalling involving female sex hormones (progesterone and estradiol). The landscape also reveals novel pathways and signalling cascades to be involved in MS, e.g. RNA binding/processing and neuronal hypoxia. Furthermore, through extensive literature searches, we were able to incorporate the mechanism of action of certain DMTs on specific molecular processes/cascades into the landscape. Currently, the action sites and mechanisms of dimethyl fumarate, fingolimod, the interferons, teriflunomide, laquinimod and natalizumab have been incorporated into the landscape in this way.

Genome-wide association studies (GWASs)
In recent years, multiple GWASs for complex genetic diseases have been conducted. In a typical GWAS, approximately 500.000 SNPs - each of which can have two variants and are collectively referred to as the disease 'genotype' - across the genome are determined (or 'genotyped') in a sample of disease cases - who all have the so-called disease 'phenotype' - and healthy controls. For each SNP, a P value is then calculated for the likelihood of a certain variant of this SNP being more prevalent in cases than controls (or the other way around). Subsequently, these SNPs - which are located in or point towards one or more genes - are then said to be associated with the disease/phenotype at a certain significance, e.g. $P < 0.05$, $P < 0.01$, $P < 0.001$ etc. In addition to the case-control GWAS, another commonly used form is the 'within case' GWAS, in which the variants of approximately 500.000 SNPs are determined in a sample that only consists of disease cases and the phenotype is disease-related, e.g. the severity of the disease or the quantitative/qualitative effect of specific drugs to treat the disease. In this study, we will perform a 'within case' GWAS together with three disease-related phenotypes, i.e. disease severity, disease burden and the effectiveness of a number of selected DMTs for MS (see below). The GWAS results will subsequently be used for gene set analysis and polygenic risk score analyses.

Gene set analysis
Gene set analysis is a recently developed method to analyse GWAS data. Through

gene set analysis, it is determined whether a whole set of genes that all belong to a certain molecular cascade or pathway is significantly associated with a given disease/phenotype. For example, a study on ADHD using gene set analysis showed significant association between all the genes belonging to a molecular landscape of ADHD that we built (Poelmans et al., 2011b) and the hyperactive/impulsive component of the disease (Bralten et al., 2013). For the gene set analysis in our study, we will divide the molecular MS landscape into a number of signalling cascades/pathways of interacting genes/proteins. Subsequently, we will ascertain whether the combined SNPs in all the genes from each cascade and from each 'within case' GWAS yield a significant P value for association with MS severity, MS burden and/or the effectiveness of selected DMTs for MS.

Polygenic risk scoring (PRS) analysis
Another recently developed method to analyse GWAS data is the PRS analysis (Euesden et al., 2015). In short, PRS analysis can be used to investigate whether there is overlap between the genotype for a certain phenotype that is usually a so-called 'trait' (e.g. the blood levels of vitamin D) and the genotype for another disease/phenotype. For this study, we will use PRS analysis to statistically quantify the genetic overlap between selected traits - based on available GWAS data - and the phenotypes of which we will conduct GWASs in this study, i.e. MS severity, MS burden and/or the effectiveness of selected DMTs for MS.
Based on our molecular landscape of MS, we will select a number of genetically determined traits to perform the PRS analyses, and we will use publicly available GWAS data for these traits, e.g. from published GWASs of vitamin D (Wang et al., 2010) and lipid/lipoprotein (Willer et al. 2013) levels. The researchers in the proposed project have recently applied PRS analysis to investigate the genetic overlap between blood levels of lipids such as cholesterol and Parkinson*s disease (Klemann et al. 2017).
Taken together, our molecular landscape of MS provides a functional map of the genes/proteins involved in MS and can be used to study the relationship between the genetic make-up of MS patients and MS-related phenotypes.

Study objective

We aim to use the results from genome-wide association studies (GWASs) of phenotypic, disease-related data of individual MS patients to provide insights into genes contributing to disease severity and burden (primary objective), and effectiveness of selected DMTs (secondary objective).

Study design

Study design: multi-center observational study, cohort design

Participants: 600 patients with MS recruited from six MS centres

Duration: 24 months

Blood sample collection and generation of genotypic (GWAS) data:
* blood sample collection: a blood sample consisting of 10 ml divided over 2 tubes will be taken in the hospital directly following a neurological out-patient visit that the patients have on a regular basis (MS patients visit the neurologist about every six months; this varies among patients); in the hospital the blood sample will be stored at -20 oC for a maximum of 3 months.
- the blood samples will be transferred to and stored at - 80oC at the Department of Molecular Animal Physiology of the Radboud University Nijmegen. Genotyping (GWAS chip: Human CoreExome-24+ DNA Analysis BeadChip, Illumina, San Diego, California) will be performed on these samples and *within-case* GWASs of the resulting genome-wide genotyping data will be conducted by Marijn Martens PhD, in collaboration with the Department of Human Genetics, Radboud University Medical Center, Nijmegen.

Phenotypic data:
* clinical data: to assess disease severity by means of the MS Severity Score (MSSS) (see below) we use the Expanded Disability Status Scale (EDSS). As the EDSS score is often used by neurologists to assess disease progression. For most patients an EDSS score will be available in the patient chart. We will ask the subset of patients for whom no EDSS score has been recorded by the neurologist to participate in a telephone EDSS scoring by a trained nurse.
* further phenotypic data will be obtained via 3 online questionnaires: MSIP, MSQoL-54 and General Patient Assessement form (questionnaires will be filled out online via the website of the MS4 Research Institute www.ms4ri.nl).

Study burden and risks

Nature and extent of the burden and risks associated with participation, benefit and group relatedness: The patients will be asked to complete three online questionnaires to gain insight into their disease severity, disease burden and disease history, as well as their use of MS medication. The patients will also be asked for a blood sample directly following a regularly scheduled visit to their neurologist. Blood sample collection will be performed by certified personnel only. A risk is that the patients are confronted with their disease (severity) by filling in the questionnaires.

Contacts

Public
Drug Target ID, Ltd.

Toernooiveld 200

Nijmegen 6500 EC
NL
Scientific
Drug Target ID, Ltd.

Toernooiveld 200
Nijmegen 6500 EC
NL

Trial sites

Listed location countries

Netherlands

Eligibility criteria

Age
Adults (18-64 years)
Elderly (65 years and older)

Inclusion criteria

600 Dutch patients with MS from six MS centres. We estimate that approximately 100 patients per centre will participate in the study, which is a feasible number of patients according to the neurologists of the respective centres.;In order to be eligible to participate in this study, subjects must meet all of the following criteria:
- diagnosed with MS for at least one year
- able and willing to participate in the study
- Dutch and from caucasian origin
- between 20 and 60 years of age;In addition, we are aiming for the gender distribution of our subjects to reflect the epidemiological findings for MS, i.e. approximately 35 % male and 65 % female subjects.

Exclusion criteria

Not applicable.

Study design

Design

Study type: Observational invasive

Masking:	Open (masking not used)
Control:	Uncontrolled
Primary purpose:	Diagnostic

Recruitment

NL

Recruitment status:	Recruitment stopped
Start date (anticipated):	01-03-2017
Enrollment:	600
Type:	Actual

Ethics review

Approved WMO Date:	14-03-2017
Application type:	First submission
Review commission:	METC Brabant (Tilburg)
Approved WMO Date:	29-03-2017
Application type:	Amendment
Review commission:	METC Brabant (Tilburg)

Study registrations

Followed up by the following (possibly more current) registration

No registrations found.

Other (possibly less up-to-date) registrations in this register

No registrations found.

In other registers

Register	ID
CCMO	NL60646.028.17