Chapter 6

Summary and conclusion
Epidemiological characteristics suggest that MS is a complex disease. Complex diseases show familial aggregation but also a high proportion of sporadic cases, no Mendelian pattern of inheritance, reduced penetrance, phenocopies, etiologic and genetic (locus and allelic) heterogeneity.

Chapter 1 provides a general introduction into MS. In Chapter 2, support for genetic influence in the etiology of MS is elaborated. Observational studies show that the recurrence risk for relatives of MS patients is increased proportional to the degree of kinship, probably as a manifestation of the amount of DNA shared. Familial clustering may also be the consequence of shared environment, but various types of epidemiological studies (adoptees, half-sibs) argue against the importance of this factor in familial clustering in MS.

The segregation of MS in multiplex families is not consistent with a fully penetrant single gene disorder. It is suspected that mutations leading to MS occur at multiple genes. The genetic model is often represented as a curve where affected individuals are those that cross a biological threshold of risk. The contributing alleles at multiple independent loci may either be rare or common, dependent upon the model applied. The best fitting model is likely to be a multifactorial model with a number of genes, including a major gene in the HLA-DRDQ region, and with one or more environmental factors. HLA is estimated to account for 10%-50% of the genetic component of MS susceptibility in Caucasian populations of northern European descent.

In order to investigate the susceptibility genes contributing to MS, different methods of genetic analysis have been applied including linkage and association analysis, Haplotype Sharing analysis and nonparametric methods like affected sibpair analysis and the transmission/disequilibrium test.

Linkage analysis assesses the segregation of a genetic variant in families with multiple affected members. In MS and other complex diseases, linkage analysis has only had limited success. The majority of positive linkages for the same disease could not be replicated in subsequent studies. No single study design produced consistently more significant results. It has been shown that for loci with a modest relative risk, that are likely to contribute to susceptibility to complex diseases like MS, the sample size needed for linkage mapping is very large.

Association studies test whether a genetic marker (polymorphism or haplotype) occurs more frequently in cases than in controls. If significant association emerges and the possible bias of population stratification is excluded, the polymorphism itself is either in the susceptibility locus or in linkage disequilibrium (LD) with the susceptibility locus. The mapping interval in association studies is typically much smaller than in linkage studies.
In monogenic diseases, single-locus association methods have proven to be powerful tools in mapping disease loci and identification of genes. However, in MS and other complex diseases single-locus methods have so far had very limited success. If indeed a complex disease is caused by a number of genes, and for any of these genes by frequent alleles with each a small effect, very large sample sizes and very dense marker sets will be necessary to find these genes by means of association analysis.

LD-based methods use multilocus association analysis. These methods are based on the assumption that affected individuals in the present generation have inherited their susceptibility alleles from common ancestors. A disease-related mutation has been introduced on a haplotype in a previous generation. With each generation, recombination events during meiosis may have led to reduction of the length of the original haplotype. Therefore, only if the distance between a marker locus and a disease locus is small, they are expected to be in LD.

Association-based studies using LD mapping are more sensitive to find also minor genes. However, also with multilocus analysis, a sufficient density of markers is necessary. The marker density needed depends on the average extent of LD and the number of haplotype blocks. On the basis of simulation studies it has been estimated that a useful level of LD is unlikely to extend beyond an average distance of about 3 kb in the general population. This implies that approximately 500,000 to 1,500,000 SNPs will be required for whole genome studies. Some authors claim, however, that in special populations LD may be spanning megabases and that much lower marker densities could generate substantial power.

As a consequence of LD, analysis of haplotype overlap can be used to map a disease gene. Once it is certain that haplotypes contain a disease gene, the smallest fragment shared by patients and not by controls is likely to contain this gene. However, when it is not certain that haplotypes contain a disease gene, the overlap of haplotypes needs to be evaluated statistically. The Haplotype Sharing Statistic (HSS) is described as a non-parametric multilocus association method for fine-mapping disease genes. HSS determines the extent of haplotype sharing for all pairs of haplotypes of patients and of controls and calculates the difference in mean haplotype sharing between patients and controls.

The underlying hypothesis of HSS is that at loci involved in the disease, haplotype sharing among patients is larger than among controls, because (i) haplotypes containing risk alleles are likely to be more often similar, especially close to the risk allele and (ii) haplotypes containing the risk alleles may be shared over longer stretches. The first factor is understandable with the
concepts of association and LD. The second factor is the consequence of the patient haplotypes containing mutations that are genetically younger than the wild type allele. Thus, fewer meioses and consequently fewer recombinations have taken place on the patient haplotypes containing a disease mutation than on the wild type haplotype.

In some models of disease, HSS is more powerful and accurate than association analysis and TDT are. Furthermore, as compared with association analysis, HSS has been shown to extract additional information from the data. However, it is unlikely that a single method would provide optimal power under all circumstances to detect susceptibility genes for complex diseases. Moreover, most often the underlying genetic model is not known. Therefore, it is recommendable to apply different methods and compare the results or even combine p-values after correction for correlation between statistics.

Chapter 3 describes the application of HSS in MS. Although an association between MS and specific HLA-types has been known for almost thirty years, the nature of this relationship has remained unclear. Furthermore, genetic resolution sufficient to implicate a specific gene in the HLA-region has not been achieved. Many loci in the HLA-region have been found significantly associated with MS, which is largely explained by the extended haplotype sharing and varying marker informativity of the region.

Using a set of 22 microsatellite markers covering the HLA-region we determined 248 haplotypes of multiple sclerosis (MS) patients from the population of the Northern Netherlands and 226 haplotypes of their relatives as controls. The data were analyzed using standard association methods and HSS. Haplotype sharing was found to be significantly greater among patients than among controls in a region of 1.1 Mb between markers G511525 and TNFa. This region was also supported by association analysis and Transmission/Disequilibrium Test (TDT). Within this region, HSS indicated the interval of 51 kb between markers G511525 and D6S1666 as the interval most likely to contain a susceptibility gene for MS. According to our present knowledge DQB1 is the sole gene in this interval. Therefore, the results of our analysis suggest that this gene plays a role in the pathogenesis of MS.

In Chapter 4, the inheritance mode of the HLA class II alleles was investigated. Haplotype clustering had indicated that two different ancestral haplotypes were likely to include a polymorphism involved in susceptibility to MS. We performed genotype association analyses for both marker loci separately and for the two-locus haplotype of markers G511525 and D6S1666. The two-locus genotype association analysis showed that in individuals who carried only one of the risk haplotypes, the risk for MS was moderately increased (OR
2.82; 95%CI 1.50 - 5.31). However, in individuals carrying two risk haplotypes the risk for MS was highly increased compared with individuals who carried no risk haplotypes (OR 37.00; 95%CI 8.31-164.74). Therefore, we concluded that this susceptibility locus for MS seems to follow an intermediate mode of inheritance. Fitting additive, multiplicative and third power risk models to the data, the effect appeared to be significantly stronger than additive.

In Chapter 5 we examined the difference between relapsing-onset and primary progressive patients. Most commonly, the disease course is characterized by relapses with worsening of symptoms followed by remissions (relapsing-remitting MS (RR-MS)). On average after 10-15 years, most RR-MS patients experience progression of symptoms between relapses and they enter the so-called secondary progressive phase (SP-MS). In a smaller number of MS patients the symptoms are progressive from onset without clear-cut relapses. This course of disease is called primary progressive MS (PP-MS).

Whether RR/SP-MS and PP-MS are two ends of a spectrum or two different diseases is not known. Apart from their different course in time, they possibly differ in clinical characteristics, such as age at onset and sex distribution, dissemination across the central nervous system, pathological, radiological, immunological and genetic characteristics. There are reports that the HLA association is different in PP-MS compared with RR-MS, but this is not a constant observation.

In order to investigate whether in our population a genetic difference in the HLA region existed between RR- and SP- versus PP-MS, we compared these two patient subgroups with respect to haplotype sharing, allelic and haplotypic association. We added five markers to the telomeric end of the studied region in order to examine whether or not we could strengthen the allelic association, haplotype sharing and transmission distortion that we had previously found in that region.

We found differences both in allele and haplotype sharing and in length of haplotype sharing between patients with RR-/SP-MS and PP-MS. In the HLA class II region containing HLA-DR and DQ, association and haplotype differences were predominant in the RR-/SP-patients. However, in a region telomeric of HLA, we found a difference in association that was almost entirely caused by PP-MS patients. This supports the hypothesis that the different types of MS are part of a spectrum and that disease-modifying genes are located within and close to the HLA region.

Returning to Chapter 2, results of candidate gene approaches are shortly discussed and the results of the many genome screens in MS more extensively. Association studies into candidate genes that show positive results will only rarely be confirmed by a second investigation. This is caused by a
number of factors, including bias and lack of study power, but also the publication selection effect of many studies on the same question and the ratio of “true relationships” to “no relationships” among those tested in the field, i.e. the prior probability of the research finding being true.

So far, only the HLA region has repeatedly shown significant association and linkage with MS. An important question is, whether the association between HLA and MS is due to a causative role of HLA-molecules in the susceptibility to MS, or whether a separate gene, not functionally involved in the HLA-system, plays a role. The HLA-system plays an important role in the immune-system. Therefore, it could indeed be functionally involved in MS, in which immune-related mechanisms likely play a role. Specific HLA-types may provide a selective evolutionary advantage from the point of view of infectious disease despite the fact that they make the individual more susceptible to an auto-immune disease. Another possibility is that a gene that is genetically close to the HLA-genes could “hitch hike” with such a haplotype, even though on itself, it might give a selective evolutionary disadvantage. In this way, also genes that contain a mutation that makes the individual susceptible to a disease, can reach higher population frequencies than expected on the basis of their selective disadvantage.

Neither whole genome screens based on linkage analysis nor screens based on association and or LD-mapping have convincingly indicated the involvement of loci other than the HLA-region. The fact that only the HLA region has been confirmed to contain a susceptibility gene for MS may well be the consequence of sample sizes and marker densities being still insufficient with regard to the complexity of the underlying genetic model. Genetic (locus and allelic) heterogeneity, reduced penetrance and contributions of environmental factors complicate the finding of susceptibility genes. It is to be expected that in the near future possibilities to find these genes will be much better with the availability of more SNP markers and of high throughput genotyping. However, funding will have to be sufficient as well, since large numbers of DNA samples will have to be examined. In order to obtain these large numbers of DNA samples, national and international cooperation will be required.