Environmental risk factors for the occurrence of multiple sclerosis

Martin Biström
To my family

“Suspicion often father of truth”

Charlie Chan at the Race Track (1936)

“A child of five could understand this.
Send someone to fetch a child of five.”

Groucho Marx
Table of Contents

Abstract........................................................................................................................................... iii
Abbreviations .................................................................................................................................... iv
Publications and manuscripts ........................................................................................................ vi
Kort sammanfattning på svenska .................................................................................................... vii

Background....................................................................................................................................... 1
  Multiple sclerosis ........................................................................................................................... 1
  Pathology ......................................................................................................................................... 2
  Diagnosis .......................................................................................................................................... 3
  Treatment ......................................................................................................................................... 3
  Etiology ............................................................................................................................................ 4
  Epidemiology ................................................................................................................................... 5
  Risk factor: Obesity ......................................................................................................................... 6
    Leptin .............................................................................................................................................. 6
    Effects of leptin on the immune system ......................................................................................... 7
    Leptin and MS ............................................................................................................................... 8
    Insulin ............................................................................................................................................ 8
  Risk factor: Herpesviruses ............................................................................................................... 9
    Epstein-Barr virus ......................................................................................................................... 10
    EBV and MS ................................................................................................................................. 10
    Human herpesvirus 6 .................................................................................................................... 11
    HHV-6 and MS .............................................................................................................................. 12
  Risk factor: Vitamin D .................................................................................................................... 12
    Effects of vitamin D on the immune system ............................................................................... 13
    Vitamin D and MS ....................................................................................................................... 14

Research rationale .......................................................................................................................... 15

Aims .................................................................................................................................................. 16

Methods .......................................................................................................................................... 17
  Study design and case ascertainment .............................................................................................. 17
  Biochemical methods ...................................................................................................................... 20
    Study I and IV ............................................................................................................................ 20
    Study II ......................................................................................................................................... 21
    Study III ....................................................................................................................................... 21
  Statistical methods ......................................................................................................................... 22
  Ethical considerations ..................................................................................................................... 24

Results ............................................................................................................................................ 26
  Characteristics of cases and controls ............................................................................................. 26
  Study I - HHV-6A and HHV-6B .................................................................................................... 26
  Study II – Leptin and Insulin .......................................................................................................... 27
  Study III – Vitamin D ..................................................................................................................... 28
  Study IV - EBV .............................................................................................................................. 30
Discussion........................................................................................................... 32
  Study I – HHV-6A and HHV-6B........................................................................... 32
  Study II – Leptin and Insulin.............................................................................. 35
  Study III – Vitamin D......................................................................................... 37
  Study IV – EBV.................................................................................................. 41
  General limitations.............................................................................................. 43

Conclusions............................................................................................................ 45

Future prospects..................................................................................................... 46

Acknowledgement ................................................................................................. 47

References ............................................................................................................... 49
Abstract

**Background.** Multiple sclerosis (MS) is an inflammatory and degenerative disease of the central nervous system that typically debuts around age 30. About 2.3 million people are affected in the world today, and besides trauma it is the most common cause of neurological disability among young adults in the western world. The disease likely develops via a complex interplay of genetic vulnerability and environmental risk factors, and adolescence is assumed to be a critical time for disease initiation. The aim of this study was to investigate how MS risk in different age groups is influenced by vitamin D, infections with Epstein-Barr virus and Human herpesviruses 6A and B as well as the metabolic markers leptin and insulin.

**Methods.** In this nested case-control study we identified pre-symptomatically drawn blood samples from individuals below age 40, that later developed relapsing remitting MS. This was done through crosslinking of the Swedish MS registry, or a local database, with six Swedish biobanks containing remainders of samples used in microbiological analyses. For each case, one control matched for biobank, sex, date of sampling and age of sampling was selected. These samples were then analysed to determine antibody reactivity against Epstein-Barr virus and Human herpesvirus 6A and B, as well as measure concentrations of leptin, insulin and 25-hydroxyvitamin D. The effect of these variables on MS risk was estimated using conditional logistic regression, both in the entire case-control material as well as stratified into three groups by age at sampling (<20, 20-29 and 30-39) and by sex.

**Results.** Human herpesvirus 6A, but not B, was consistently associated with an increased risk of developing MS. In contrast, Epstein-Barr virus demonstrated an age dependent pattern indicating that early infection may be protective against MS while later infection increases the risk. As for the metabolic markers, insulin was not associated with MS while elevated levels of leptin showed an association with increased MS risk both among individuals below 20 years of age and among all men. For women there was instead an inverse association in the oldest group, aged 30-39, when adjusting the leptin analysis for insulin concentrations. Finally, having vitamin D concentrations in the top quintile was associated with decreased MS risk, without evidence of a stronger effect in young subjects.

**Conclusion.** These results implicate Human herpesvirus 6A and leptin as risk factors for MS development. They also further support a protective role for vitamin D in MS etiology and provide serological evidence of an age dependency of Epstein-Barr virus infection as it relates to MS risk.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>25(OH)D</td>
<td>25-hydroxyvitamin D</td>
</tr>
<tr>
<td>AA</td>
<td>Amino Acid</td>
</tr>
<tr>
<td>AP</td>
<td>Attributable Proportion</td>
</tr>
<tr>
<td>BMI</td>
<td>Body-Mass Index</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>CSF</td>
<td>Cerebrospinal Fluid</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>DMT</td>
<td>Disease Modifying Treatment</td>
</tr>
<tr>
<td>EAE</td>
<td>Experimental Autoimmune Encephalomyelitis</td>
</tr>
<tr>
<td>EBNA</td>
<td>Epstein-Barr Nuclear Antigen</td>
</tr>
<tr>
<td>EIMS</td>
<td>Epidemiological Investigation of Multiple Sclerosis</td>
</tr>
<tr>
<td>GEMS</td>
<td>Genes and Environment in Multiple Sclerosis</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome Wide Association Study</td>
</tr>
<tr>
<td>HHV-6A</td>
<td>Human Herpesvirus 6A</td>
</tr>
<tr>
<td>HHV-6B</td>
<td>Human Herpesvirus 6B</td>
</tr>
<tr>
<td>HLA</td>
<td>Human Leukocyte Antigen</td>
</tr>
<tr>
<td>IE1</td>
<td>Immediate Early Protein 1</td>
</tr>
<tr>
<td>IM</td>
<td>Infectious Mononucleosis</td>
</tr>
</tbody>
</table>
IMSE  Immunomodulation and Multiple Sclerosis Epidemiology

LC-MS/MS  Liquid Chromatography Tandem Mass Spectrometry

MFI  Median Fluorescence Intensity

MHC  Major Histocompatibility Complex

MRI  Magnetic Resonance Imaging

MS  Multiple Sclerosis

OR  Odds Ratio

PCR  Polymerase Chain Reaction

PHAS  Public Health Agency of Sweden (Folkhälsomyndigheten, fd Smittskyddsintitutet)

PPMS  Primary Progressive Multiple Sclerosis

RRMS  Relapsing Remitting Multiple Sclerosis

SPASM/STOPMS  Stockholm Prospective Assessment of Multiple Sclerosis

SPMS  Secondary Progressive Multiple Sclerosis

UV  Ultraviolet

VCA  Viral Capsid Antigen
Publications and manuscripts


Kort sammanfattning på svenska

**Bakgrund.** Multipel skleros (MS) är en inflammatorisk och degenerativ sjukdom i det centrala nervsystemet som oftast drabbar personer i åldern 20–40 år. Cirka 2,3 miljoner människor världen över har MS och bortsett från trauma så är det den vanligaste orsaken till neurologisk funktionsnedsättning bland unga vuxna ivästvärlden. Symtomen varierar beroende på vilken del av nervsystemet som påverkats och kan omfatta sådant som störd känsel, paraser, trötthet, smärta, balansrubningar och autonom dysfunktion. Orsaken till sjukdomen är inte klarlagd men den rådande hypotesen är att den beror på ett komplext samspel mellan genetisk sårbarhet och faktorer i miljön såsom infektion med Epstein-Barr virus, låga nivåer av D-vitamin, rökning och övervikt i ungdomen. Syftet med den här studien var att undersöka hur tre av dessa faktorer, samt infektion med Humant herpesvirus 6A och B, påverkar risken att utveckla MS.

**Metod.** Genom samkörning av det svenska MS registret, eller en lokal MS-diagnosdatabas i Umeå, med sex svenska biobanker identifierades blodprover från personer som utvecklat MS. Dessa biobanker är kopplade till mikrobiologiska laboratorier vid universitetssjukhus i Umeå, Skåne, Göteborg, Örebro och Linköping samt Folkhälsomyndigheten. Inklusionskriterier för studien var MS med skovvist förlopp, att provet som identifierats var taget innan symptomdebut samt att provet var taget före 40 års ålder. För varje sådant fallprov valdes sedan en kontroll ut. Dessa kontroller matchades för biobank, kön, protagningsdatum samt protagningsålder. Proverna analyserades med avseende på virusantikroppar, D-vitamin samt de två hormonen leptin och insulin som är kopplade till både övervikt och metabol funktion. Riskkvoeter för de olika exponeringarna beräknades med logistisk regression, både i materialet som helhet och stratifierat utifrån ålder och kön.

**Resultat.** Höga nivåer av antikroppar mot Humant herpesvirus 6A, men inte B, var konsekvent associerat med förhöjd MS-risk. Förekomst av antikroppar mot Epstein-Barr virus uppväside en åldersberoende association till MS där infektion före 20 års ålder var en skyddande faktor medan infektion senare i livet istället var associerad med ökad risk för MS. Höga nivåer av D-vitamin var associerat med minskad MS-risk. För insulin sågs ingen association till MS men höga nivåer av leptin var förknippad med förhöjd MS-risk före 20 års ålder samt i gruppen män.

**Tolkning.** Dessa resultat indikerar att infektion med Humant herpesvirus 6A samt födröjd infektion med Epstein-Barr virus kan öka risken för att utveckla MS. Fynden av en koppling mellan MS-risk och nivåer i blodet av leptin respektive D-vitamin stärker dessutom tidigare antaganden om att låga nivåer av D-vitamin samt övervikt tidigt i livet är riskfaktorer för MS.
Background

Multiple sclerosis

Multiple sclerosis (MS) is an inflammatory and degenerative disease of the central nervous system (CNS) afflicting approximately 2.3 million individuals worldwide.\(^1\) Its incidence peaks between 20-40 years of age and besides trauma it is the most common cause of neurological disability among young adults in the western world. Signs and symptoms depend on what part of the CNS that is affected by the disease and can include a wide variety, e.g. sensory disturbances, paresis, ataxia, pain and autonomic dysfunction. The most common presentation is a relapsing remitting form of MS (RRMS) where neurological deficits appear sub acutely over days or weeks and then gradually disappear over a longer period, resulting in either partial or complete clinical recovery. Left untreated, most individuals will eventually progress to a phase of more continual worsening known as secondary progressive MS (SPMS). The progressive form is characterized by neurodegeneration with steadily increasing neurological deficits with or without superimposed relapses.

In contrast to the relapsing form of MS about 5-15% of individuals who are diagnosed will have a disease course that is progressive from the start, so called primary progressive MS (PPMS), the most common initial presentation of which is spastic paraparesis. While RRMS has a female predominance - being about 2 or 3 times as common among females compared to males - no such clear difference is seen regarding PPMS. Although it has by some been considered a separate entity from relapsing MS, the similarities between SPMS and PPMS are striking and the division of MS into different subtypes may have been driven by drug licencing concerns.\(^2\) PPMS usually debuts about 10 years later than RRMS and both forms of progressive MS have a similar median age at onset of about 40 years.\(^3\)

As evidenced by the numerous studies of the natural history of multiple sclerosis conducted before the introduction of the first disease modifying treatments (DMTs) in the 1990s, MS is a heterogenous condition that often has severe consequences if left untreated. Without treatment individuals with MS has a 5-10 years shorter life expectancy compared to the general population. The median time from disease onset until the need for a walking aid is 20 years and the time
to reach such disability milestones is not obviously affected by the number of relapses.\textsuperscript{3,4}

**Pathology**

As the name multiple sclerosis (many scars) suggests, MS is defined by demyelinated plaques disseminated throughout the CNS, that with time are converted to glial scars by astrocytes.\textsuperscript{5} Active plaques contain large amounts of activated macrophages\textsuperscript{6} as well as infiltrates of lymphocytes, dominated by T-lymphocytes, mainly CD8+ T-cells, with B-lymphocytes and plasma cells present to a lesser extent. Early on, this inflammatory process selectively targets myelin and oligodendrocytes while partly sparing neurons, enabling some plaques to be repaired by remyelination. Later stages are however characterised by axonal loss and associated irreversible neurological deficits.\textsuperscript{5} Anatomical sites in the CNS where MS-lesions commonly are located include juxtacortical and periventricular white matter, optic nerves as well as infratentorial sites such as the brainstem, cerebellum and spinal cord. Just as MS is an heterogenous disease in a clinical sense, there are reports of pathological heterogeneity as well, indicating that disease mechanisms may differ from patient to patient.\textsuperscript{6}

While there are no clear-cut pathological distinctions to be made between RRMS and SPMS/PPMS, as both clinical phenotypes contain elements of inflammation, the inflammatory infiltrates appear to contain a higher proportion of plasma cells and B-cells in progressive disease. As a general rule, relapsing MS more commonly has active lesions with lymphocytic inflammation whereas progressive disease more often have lesions with an inactive core surrounded by activated macrophages and microglia.\textsuperscript{2} Additionally, progressive MS is also characterized by subpial cortical lesions\textsuperscript{7} associated with tertiary lymphatic tissue found in the meninges.\textsuperscript{8} What clearly differentiates MS pathology from other diseases of the CNS is the pattern where MS-lesions are selectively perivenous (i.e. located in close proximity to veins and venules) and primarily demyelinating with oligodendrocyte loss. Besides these inflammatory lesions in white and grey matter, there is also a component of neurodegeneration of the brain and spinal cord in a more general sense, leading to a CNS atrophy that is more pronounced in the progressive phase of the disease. This neurodegeneration occurs also in the earliest phases of MS but is amplified by the accumulated lesion burden and brain aging associated with progression. Both the degeneration and the demyelination are believed to be mediated by mitochondrial and oxidative injury, resulting in downstream consequences such as cellular energy deficiency and ionic
imbalance. The root cause of this destructive cascade is still unresolved, however.7

**Diagnosis**

Along with advances in technical development, most importantly magnetic resonance imaging (MRI), the diagnostic criteria for MS are continually updated to allow earlier and more specific diagnoses, with the latest revision published in 2017.9 Thankfully, the delay from onset of symptoms until diagnosis has steadily decreased from the 1980s to the early 2000s.10 An early and accurate diagnosis has become increasingly important in the last 25 years as the arsenal of DMT options have increased dramatically11 and evidence is accumulating that early effective treatment is important to reduce disability.10 Since there is no pathognomonic clinical feature or test for MS, the diagnosis relies on combining findings from all available sources such as imaging, history, neurological examination and laboratory analyses.

The diagnosis of multiple sclerosis is based around the ability to demonstrate dissemination in space (distinct anatomical locations) and time (development over time) of CNS lesions in an individual presenting with a typical clinically isolated syndrome. Such a syndrome is defined as an episode of symptoms and objective findings typical of a multiple sclerosis relapse, lasting at least 24h, that reflect one or more demyelinating lesions in the CNS, in the absence of fever or infection. Dissemination in space can be determined by MRI if at least one lesion is identified in at least two of four typical anatomical sites (juxtacortical/cortical, periventricular, infratentorial or spinal) or by clinical attacks implicating separate CNS sites. Dissemination in time can also be determined through MRI, either by detection of new lesions on separate examinations or by detection of both contrast enhancing and non-enhancing lesions at the same examination. Additional ways this criterion can be fulfilled is through an additional clinical attack or the presence of oligoclonal bands in the cerebrospinal fluid (CSF).9

**Treatment**

Treatments can be either disease modifying or symptomatic and both categories are important for managing MS. Especially the disease modifying variety has seen incredible progress since the introduction of interferon-β back in the mid-1990s.
This treatment reduced relapses with about 30% while also showing modest effect on slowing disability progression. With the introduction of natalizumab in 2006, the first monoclonal antibody licensed for use in MS, a new era of more efficacious treatments began and there are many different substances available today. Unfortunately, there is still no cure for MS however, and DMTs are prone to cause adverse effects, often relative to their effectiveness, making the disease problematic both for afflicted individuals as well as society as a whole.

**Etiology**

While the exact cause of this debilitating disease still is unknown, many risk factors have been identified and it is largely considered to be a consequence of a complex interplay between genetic susceptibility and environmental exposures. Twin studies have shown concordance rates among monozygotic twins of about 25%, driven largely by female twins, which makes it clear that factors other than genes are involved in disease development. In recent years large genome wide association studies (GWAS) have implicated over 200 risk loci, most of which are proposed to influence the immune system. The strongest genetic risk factors are common variants of human leukocyte antigens (HLA) located on chromosome 6. The two single nucleotide polymorphisms with the strongest association with MS are the presence of HLA-DRB1*15:01 and absence of HLA-A*02:01, with a combined fivefold increase in risk.

Environmental risk factors currently considered to have strong evidence linking them to the risk of developing MS are Epstein-Barr virus infection, smoking, low serum levels of vitamin D and obesity during adolescence. Since MS is a rare disease insofar as the incidence is relatively low, it is impractical (and often unethical) to use the gold standard method for determining cause and effect in medicine (i.e. randomized controlled trials) to test if these associations are causal. Instead the evidence comes from mainly three different sources, observational epidemiology in the form of cohort studies and case-control studies, Mendelian randomization studies and experimental studies utilizing experimental autoimmune encephalomyelitis (EAE), the animal model of MS.

In EAE an inflammatory response against CNS antigens is induced, mainly through transfer of pathogenic T cells or by immunisation with suitable peptides. The most common species are mice but rats, primates and recently also zebrafish.
have been used as well. While it is clear that EAE has contributed to the development of numerous DMTs there are also instances where the model has led researchers astray. One example is glatiramer acetate that was originally designed to induce EAE by mimicking myelin basic protein but is now instead used as a treatment of relapsing MS. Another is example is that inhibition of tumor necrosis factor, which showed improvements in EAE, turned out worsen MS symptoms when tested in clinical trials.

Despite the lack of an identified autoantigen in MS, the currently dominating theory of disease mechanisms underlying disease development is an autoimmune reaction directed against the CNS. Supporting this claim of autoimmunity as the primary driver of tissue injury are the many similarities between MS and other autoimmune diseases. One such similarity is that both genetic and environmental factors are implicated in the etiology and that most of the genes associated with MS risk have effects on the immune system. Studies investigating whether patients with MS and their first-degree relatives has an increased risk of other autoimmune diseases have been conflicting, however. While there are some indications of increased risk of thyroid disease among both MS patients and their relatives, as well as a slightly increased risk of inflammatory bowel disease and psoriasis among individuals with MS, it has been suggested that such findings in large part may be explained by surveillance bias due to increased contact with health care after MS diagnosis.

**Epidemiology**

One of the most striking features of MS is the fact that the disease becomes more common (i.e. has a higher prevalence) further away from the equator. This phenomenon is sometimes referred to as a latitude gradient and has inspired research efforts focusing on explaining this finding. There are many factors that could explain such patterns, one being genetics. Genetic differences are unlikely to the entire explanation however, as the gradient persists even after adjusting for HLA-DRB1 allele frequency. Other explanations that have been put forward are sun exposure, vitamin D levels, and differences in the panorama or timing of infections. Support for the importance of sun exposure comes from an elegant study utilizing NASA satellite data to show that available ultraviolet (UV) radiation is at least 20 fold more predictive of MS prevalence than latitude.
Additional epidemiological evidence for the involvement of environmental factors in MS etiology comes from studies of disease occurrence among migrants. Individuals that move from a high-risk area (i.e., a country where the disease is common) to a more low-risk area, appears to have a lower disease risk, and studies of the effect of age at migration indicates that the MS risk of an individual is determined early in life.\textsuperscript{22} The opposite effect, meaning a risk increase among individuals moving from a low-risk to a high-risk area has been more difficult to establish, but has been confirmed in more recent studies performed in Scandinavia.\textsuperscript{23,24} Especially interesting was the finding of higher risk among immigrants arriving before 15 years of age, lending further support to the hypothesis that the risk of developing MS is influenced by one or more environmental factors early in life.\textsuperscript{24}

**Risk factor: Obesity**

Overweight, defined as a body-mass index (BMI) of 25 kg/m\textsuperscript{2} or greater, and obesity (≥30 kg/m\textsuperscript{2}) is increasing worldwide\textsuperscript{25} and induces a state of low-grade inflammation due to the nutrient and energy overload associated with accumulation of excess white adipose tissue. This inflammatory state is chronic in nature, of medium to low intensity and is characterized by locally elevated levels of inflammatory cytokines as well as immune cell infiltration of metabolic tissues.\textsuperscript{26} In the last decade obesity, especially when present during adolescence, has become recognized as a risk factor for MS with data supporting this connection coming from both cohort,\textsuperscript{27,28} case-control,\textsuperscript{29–32} and Mendelian randomization studies.\textsuperscript{33,34} There are also data suggesting that obesity interacts synergistically with genetic risk factors\textsuperscript{35} as well as with infectious mononucleosis,\textsuperscript{36} resulting in a many times higher MS susceptibility if present in combination with either. While most studies linking adolescent obesity to an increased MS risk have used either BMI or body silhouettes to estimate levels of adiposity, advances in the understanding of white adipose tissue as an endocrine organ and the emerging concept of metabolic syndrome implicates other possibilities as well.

**Leptin**

Maintenance of adequate energy reserves is important for the survival of any organism, and in humans these stores are mostly made up of triglycerides inside specialised fat storing cells (i.e. adipocytes). These cells, besides acting as energy deposits containing triglycerides, also produce signalling molecules that are introduced into systemic circulation (i.e. hormones) commonly known as
adipokines or adipocytokines, one of the most well studied to date is leptin. The name leptin originates from the Greek word leptos, meaning “thin” and was discovered in 1994.\textsuperscript{37} It was revealed to be the peptide product of the mouse recessive obese (ob) allele, responsible for body weight homeostasis, the lack of which was what caused ob/ob mice to become not just obese but also diabetic, hyperphagic and sterile.\textsuperscript{38} The discovery of leptin as a way for an organism to sense energy stores was an important step in unravelling the mystery of how an almost constant bodyweight can be maintained year after year without counting calories,\textsuperscript{39} but the initial hopes of treating obesity by manipulating the leptin axis has not yet been fulfilled.

While adipocytes located in white adipose tissue mainly produce leptin in proportion to the magnitude of their triglyceride content, there are many other mechanisms regulating its secretion as well as a few other known tissues where production takes place. Among variables that increase serum leptin levels are insulin, glucocorticoids, tumor necrosis factor alpha, oestrogen, Interleukin-1 and alcohol while androgens, growth hormone and cigarette smoking are reported to decrease leptin levels.\textsuperscript{40} Despite a multitude of factors influencing leptin concentrations it has been shown to correlate well to fat mass\textsuperscript{41} and is therefore a suitable serum marker to estimate adiposity.

Although one of its functions is fat mass homeostasis through the binding of the leptin receptor in the hypothalamus, resulting in a reduction of calorie intake and an increase in energy expenditure, leptin is not a satiety factor as it does not increase acutely post prandially.\textsuperscript{42} It is instead released into the bloodstream in a pulsatile manner, with diurnal variations resulting in concentrations peaking during the night.\textsuperscript{43} The half-life of the bound form of leptin is just over one hour while the free form is cleared faster with a relatively short half-life of about 3 minutes, similar to other peptide hormones such as insulin.\textsuperscript{40}

**Effects of leptin on the immune system**

Besides well documented effects on appetite regulation, energy expenditure and reproductive function, leptin is now also recognized as an important link between adipose tissue and the immune compartment. This includes both cells belonging to the innate as well as the adaptive branch of the immune system, as most immune cell types express the leptin receptor. Leptin is currently considered to
be proinflammatory and may contribute to the low-grade inflammation seen in overweight individuals. It has also been proposed as a link between adiposity and some obesity-associated diseases involving excessive or inappropriate inflammation, including autoimmune diseases. More specifically, as it relates to lymphocytes that have been implicated in MS pathogenesis, leptin promotes proliferation of both naïve B and T cells and increases CD4+ helper T (T\textsubscript{h}) polarization towards the proinflammatory T\textsubscript{h}1 phenotype. It also decreases proliferation of the immune suppressing regulatory T (T\textsubscript{Reg}) cells while increasing proliferation and function of T\textsubscript{h}17 phenotypes. Additionally, leptin may reduce B cell apoptotic rate, modulate their development and increase B cell cytokine production.

**Leptin and MS**
Evidence for direct associations between leptin and MS is sparse, but there are some indications of a role for leptin in MS etiology. For one, leptin has been shown to be indispensable for the induction of EAE, and a surge in serum leptin seems to precede the clinical onset. Human case-control studies of serum leptin concentrations in individuals with established MS compared to controls has shown conflicting results. However, according to a meta-analysis combining data from nine studies there are indications that MS patients have higher leptin levels.

**Insulin**
Insulin, a peptide hormone produced by β cells in the pancreas, was discovered back in the early 1920ies. This event was deemed so important that it was awarded with a Nobel prize in 1923 and exogenous insulin was quickly put to clinical use in the treatment of diabetic children, with dramatic lifesaving effects. Today insulin is used to also treat type 2 diabetes mellitus, a disease defined by hyperglycaemia caused by relative insulin insufficiency, due to the development of resistance in certain tissues to the effects of insulin. Generally speaking, the role of insulin in glucose homeostasis is primarily mediated by its anabolic effects on the liver, skeletal muscle and white adipose tissue. While insulin resistance is incredibly complex and still not fully understood, it is clear that it is associated with obesity. Most obese insulin resistant individuals do not go on to develop type 2 diabetes however, as the pancreas more often than not is able to compensate for this resistance by producing even more insulin, and thereby maintain glucose tolerance. This may lead to a state of hyperinsulinemia, characterized by chronically elevated insulin levels without
concomitant hypoglycaemia. Complicating matters further however, are observations that hypersecretion of insulin and hyperinsulinemia are more common in obesity than insulin resistance is, creating uncertainties about which way the arrows of causality are pointing. Individuals with established MS have been shown to have a high prevalence of insulin resistance but there is currently no evidence for a direct involvement of insulin or insulin resistance in MS etiology.

**Risk factor: Herpesviruses**

Ever since multiple sclerosis was recognised as a specific disease during the 19th century, infections have been suggested as being the possible cause. The enthusiasm for this line of reasoning may have many sources, one being the success in finding viruses responsible for other demyelinating diseases such as subacute sclerosing panencephalitis and in more recent times progressive multifocal leukoencephalopathy. The list of infectious agents that have been considered in MS etiology is quite long, mostly dominated by viruses, but most of these hypotheses have been abandoned along the way. An exception to this rule is the *Herpesviridae*, a virus family of which nine members are known to infect humans. These human herpesviruses are further divided into three subfamilies estimated to have arisen 180-220 million years ago: alpha (herpes simplex viruses 1 and 2, varicella zoster virus) beta (cytomegalovirus, human herpes virus 6 A and B, human herpesvirus 7) and gamma viruses (Epstein-Barr virus and human herpesvirus 8). Among these viruses, the prime candidate when it comes to evidence linking it to MS is the Epstein-Barr virus (EBV) followed by the Human herpesvirus 6A (HHV-6A).

These human herpesviruses share a basic structure including four main components, core, capsid, tegument and envelope. The core consists of a large linear double-stranded DNA that is tightly packed inside the icosahedral capsid which in turn is surrounded by the tegument, a layer of viral proteins connecting the capsid to the envelope. The envelope includes viral glycoproteins incorporated into a lipid bilayer derived from the host cell. During infection of new cells, viral glycoproteins interacts with the components of the host cell membrane, resulting in either a fusion or endocytosis and consequently a relocation of the viral nucleocapsid to the inside of the plasma membrane. After primary infection herpesviruses establish long-term latency in their respective target cells, enabling them to later reactivate and spread to other hosts.
Epstein-Barr virus
EBV was discovered in 1964 through electron microscopy examination of cultured lymphoma cells.\textsuperscript{62} While it was immediately recognized as a herpesvirus it seemed unexpectedly inert and did not appear to destroy the cell culture even when actively producing new virus particles. In fact, it later turned out that EBV had the ability to establish latent, persistent, infection \textit{in vitro} by driving B lymphoblast growth, a process known as immortalization. EBV has since then also been extensively studied \textit{in vivo} and have been found to be B lymphotropic and persist for life in resting memory B cells.\textsuperscript{63}

It has been estimated that approximately 90\% of individuals are infected with EBV before 30 years of age, as measured by seropositivity to viral antigens.\textsuperscript{64} The age at which infection occurs varies with geographical regions where almost all individuals are infected in early childhood in developing countries while developed countries show a bimodal pattern where infection peaks both in early childhood and during adolescence.\textsuperscript{65} While most individuals that are infected with EBV are asymptomatic lifelong carriers, there are disease states, aside from MS, that has been linked to presence of the virus. It was the first example of an oncogenic virus in humans with examples of virus-associated malignancies such as Hodgkin lymphoma, Burkitt lymphoma and nasopharyngeal carcinoma.\textsuperscript{66} Age at primary infection with the virus is an important aspect in determining how severe symptoms the infection will elicit. In children it rarely gives rise to serious disease but during adolescence as many as 28-75\% have been reported to develop infectious mononucleosis (IM), a disease characterized by fever, pharyngitis and lymphadenopathy.\textsuperscript{64,67} These symptoms appear not to be associated with lytic infection of B lymphocytes but rather the powerful activation and expansion of CD8 T cells and the resulting production of cytokines. Why IM is a disease of adolescents and young adults in contrast to young children is presently not well understood.\textsuperscript{67}

EBV and MS
A connection between MS and EBV was suggested as early as 1981\textsuperscript{68} based on similarities between IM and MS epidemiology. This observation has since been confirmed in studies showing that a history of IM increases the risk of developing MS approximately 2.2 fold.\textsuperscript{69} Furthermore, several nested case-control studies, using prospectively drawn blood samples, has revealed high concentrations of
antibodies against EBV nuclear antigen 1 (EBNA-1) to be a strong predictor of future risk to develop MS. In a seminal case-control study by Levin et al from 2010 that utilized a series of blood samples from a large number of individuals, the investigators showed that all persons who developed MS, but were negative to EBV as determined by the earliest available sample, seroconverted before disease onset. Serological support for or against the hypothesis that EBV infection during or after adolescence increases the risk of MS while infection during childhood may be protective is lacking however. The main reason for this lack of knowledge is the difficulty in attaining serum samples from young individuals that go on to develop MS later in life.

**Human herpesvirus 6**
The virus that in the end became known as HHV-6 was first isolated back in 1986 and at that time believed to primarily infect B-lymphocytes, hence its first name human B lymphotropic virus (HBLV). It was later recognized to infect T-lymphocytes and was subclassified into two subtypes, HHV-6A and B. These were at first considered variations of the same species, and was only quite recently recognized as two separate viruses based on differences in among other things cell tropism, DNA sequences and epidemiology. HHV-6B is recognized as the primary causative agent of exanthem subitum, a disease characterised by high fever and subsequent development of a rash commonly seen in infants. No disease has in such a clear way been associated with HHV-6A.

As indicated by the fact that the two viruses first were considered to be variants of the same virus, they share many similarities. They have an overall nucleotide sequence identity of about 90%, with some regions showing more diversity and others less. This similarity has made it difficult to separate immunological response against one from the other and many earlier publications fail to differentiate between the two viruses. This is one reason that much is still unknown about the epidemiology and disease associations of these viruses, other reasons being the ubiquity and chronicity of infection and the curious fact that HHV-6A and B, as the only human herpesviruses, are capable of integration into the human genome. It is estimated that about 1% of the population have inherited this condition of chromosomally integrated HHV-6 (ciHHV-6) which may confound studies of disease associations relying on polymerase chain reaction (PCR), as an unexpectedly high number of viral DNA copies may be detected without viral replication actually taking place.
**HHV-6 and MS**

One line of evidence linking HHV-6 to MS are studies indicating that MS brains compared to control brains have greater levels of viral DNA and mRNA, with the highest levels found in demyelinated plaques. Even more incriminating are the findings of active viral replication in oligodendrocytes, the cell responsible for production and maintenance of myelin in the CNS. Another line of evidence comes from findings of intrathecal antibodies with specificity for HHV-6 in a subset of MS patients, including in the form of oligoclonal bands. It is well known that in many infectious diseases of the CNS, oligoclonal bands with specificity against the causative agent may develop and this finding therefore has been taken to implicate HHV-6 in MS pathogenesis, at least in a subset of cases.

As for studies on immunological response against HHV-6 in serum, one study using prospective samples found higher antibody reactivity among individuals that later developed MS than among matched controls. The results for established MS have been mixed where some studies have found significant differences between cases and controls while other have not.

Looking at the two viruses, HHV-6A and HHV-6B, in relation to a disease of the central nervous system such as MS it is likely that HHV-6A is the more promising candidate because of its proposed greater neurotropism and greater *in vitro* ability to infect both astrocytes and oligodendrocytes. Further supporting this reasoning are studies that have found HHV-6A to be more prevalent in MS patients compared to controls. Substantial serological support from studies able to differ between the two viruses has been lacking, however.

**Risk factor: Vitamin D**

Around the year 1920, a fat-soluble factor with properties enabling it to cure the bone disease rickets was discovered in milk and cod-liver oil. At about the same time, an alternative cure was discovered - irradiation of the skin with UV light from quartz-mercury lamps. It was soon realized that these two separate cures were linked, in part by the surprising finding that irradiation of some foods provided them with antirachitic properties. This antirachitic fat-soluble factor was later named Vitamin D. It comes in two forms, cholecalciferol or vitamin D₃ (C₂₇H₄₄O) and ergocalciferol or vitamin D₂ (C₂₈H₄₄O). They have comparable effects on human physiology due to their highly similar structure (Figure 1), the only difference being that vitamin D₂ has an additional double bond and an extra
methyl group, but there are reports that vitamin D₃ supplementation is slightly more effective at raising an individual’s vitamin D status.⁸⁶

Figure 1. Chemical structure of vitamin D₃ (left) and vitamin D₂ (right). Source: https://pubchem.ncbi.nlm.nih.gov/compound/5280795#section=2D-Structure https://pubchem.ncbi.nlm.nih.gov/compound/5280793#section=2D-Structure

Vitamin D₃ can be synthetized from cutaneous 7-dehydrocholesterol upon exposure to UVB (wavelength 290-315 nm) and either vitamin D₂ (from ergosterol) or D₃ can be absorbed from the diet. Upon entering the circulation, it is transported to the liver where it is metabolized to 25-hydroxyvitamin D (25[OH]D), which is the metabolite that usually is measured when determining an individual’s vitamin D status. The active form of vitamin D, 1,25(OH)₂D, requires a strictly regulated hydroxylation step after which it may enter the target cell and exert its effects. The kidneys are the main site of vitamin D activation, although many different cells has this capability, a process influenced by levels of phosphorus, calcium and parathyroid hormone.⁸⁷

Effects of vitamin D on the immune system
Besides the well-known attributes of promoting bone health and regulating calcium and phosphorus homeostasis, vitamin D also has important effects on the immune system. These effects include modulation of CD4+ Tₕ and T⁵₆⁷ differentation, resulting in enhanced T⁵₆⁷ immune suppression and reduction in Tₕ1 and Tₕ17 phenotypes.⁸⁸ Both Tₕ1 and Tₕ17 cells have been implicated as key
players in MS pathogenesis\textsuperscript{89} and T_{Reg} are believed to be important for protection against autoimmunity.\textsuperscript{90} Vitamin D also reduces B cell function, by inhibiting differentiation into memory and plasma B cells as well as reducing antibody production.\textsuperscript{88} Taken together, these effects provide a theoretical mechanism by which vitamin D may affect the risk of developing MS.

**Vitamin D and MS**

At first inspired by the intriguing MS epidemiology displaying a latitude gradient where the disease is more prevalent further away from the equator, the vitamin D hypothesis of MS etiology has in recent decades gained support from many other sources. In 2004, using data from the Nurses’ Health Study I and II, Munger et al showed that taking a vitamin D supplement of $\geq$400 international units per day was associated to reduced risk of developing MS.\textsuperscript{91} Additional evidence comes from studies showing a protective effect from higher consumption of vitamin D containing foods or supplements such as cod liver oil\textsuperscript{92} and fatty fish.\textsuperscript{93}

There are also three separate studies analysing concentrations of 25(OH)D in biobanked serum samples collected before disease onset, all of which have found lower risk of MS associated with higher 25(OH)D concentrations.\textsuperscript{94–96} These studies run the risk of being confounded by sun exposure however, and evidence is accumulating that being exposed to UV radiation has immunomodulatory properties independent of vitamin D.\textsuperscript{97} Evidence strengthening the case for a causative role for vitamin D in MS etiology and disentangling this effect from UV radiation has come in recent years in the form of Mendelian randomization studies. These studies have shown that genes affecting vitamin D levels also are associated with MS risk.\textsuperscript{98,99}
Research rationale

Despite an ever-growing body of knowledge regarding the etiopathogenesis of MS, the exact cause is still unknown. While there is robust evidence linking EBV/IM, low vitamin D and overweight early in life to increased MS risk, much is yet to be understood about the mechanisms. This is even more apparent concerning metabolic markers such as leptin and insulin, as well as the two viruses HHV-6A and HHV-6B which are especially understudied in the period preceding MS onset. There are indications that adolescence is a particularly vulnerable period, where environmental factors may interact to trigger disease in genetically susceptible individuals. Environmental exposures are notoriously difficult to study at this critical timepoint however, mainly due to the delay from disease initiation until symptom onset and the fact that MS is a relatively rare condition.

Earlier studies that have been successful in gathering data on exposures during adolescence have often done so through retrospective methods such as questionnaires. These run the risk of being influenced by recall or surveillance biases. Case-control studies “nested” inside other studies that have collected exposure data for other purposes has proven to be a viable alternative. However, they have generally been lacking in statistical power to detect age specific effects due to the low number of young individuals included. By utilizing Swedish biobanks containing blood samples drawn at an early age, this thesis attempts to further the understanding of how the environment during adolescence may influence the risk of developing MS.
Aims

To investigate how environmental exposures affects the risk of developing MS, with particular focus on the time period presumed to be of greatest importance for the pathophysiological onset of the disease.

Study specific aims:

Study I: To test the hypothesis that antibodies against Human herpesvirus 6A and B are associated with increased multiple sclerosis risk.

Study II: To test the hypothesis that leptin or insulin are associated with multiple sclerosis risk.

Study III: To test the hypothesis that high levels of vitamin D are protective against multiple sclerosis and that this effect is more pronounced in young individuals.

Study IV: To test the hypothesis that infection with Epstein-Barr virus is an age-dependent risk factor for multiple sclerosis and investigate if there is an interaction between infections with Epstein-Barr virus and Human herpesvirus 6A as they relate to the risk of multiple sclerosis.
Methods

Study design and case ascertainment

All four studies are either completely or partially (study I) based on a national pre-MS case-control study of risk factors for multiple sclerosis, using presymptomatically drawn plasma or serum samples stored in Swedish biobanks. These samples were remainders after analyses performed in routine clinical practise at the Public Health Agency of Sweden (PHAS) or microbiological laboratories associated with university hospitals in Umeå, Örebro, Linköping, Skåne or Gothenburg. In order to identify available samples from individuals with MS, these six biobanks were crosslinked with either of two data sources, the Swedish MS registry (as of Feb 2012) or in the case of the Umeå biobank, a local database of MS and possible MS diagnoses in Umeå. The Swedish MS registry (now a part of the Swedish Neuro Registry, www.neuroreg.se) is a government funded national quality registry designed to improve quality of care for MS patients, that covers a majority of all prevalent MS cases in Sweden.100

The local Umeå database was created in 2009, for use in an earlier project,101 by searching medical records and national registries for MS and adjacent diagnoses. The reasoning behind the creation of this local database was the, at that point in time, low coverage of northern Sweden in the Swedish MS registry. MS diagnoses for the database were ascertained by review of medical records, using the prevailing diagnostic criteria for MS, and the date of onset was determined as the date of first symptoms suggestive of MS.

After linking the Umeå biobank to the local database, all individuals that did not fulfil the inclusion criteria or had samples in a biobank used in an earlier MS study (designated “Other subcohort” in Figure 2) were excluded. The other five biobanks were crosslinked with the Swedish MS registry to identify cases eligible for inclusion. In the next step, all were excluded that failed to meet inclusion criteria based on data from the registry. This was followed by additional exclusions, mainly due to samples missing in freezers but also as a consequence of new data emerging as missing registry data was completed by review of medical records. When one person had samples in multiple biobanks the oldest available sample was chosen, and the rest excluded (designated “Duplicate” in Figure 2). In addition to this, a total of 10 individuals (5 cases and 5 controls) declined participation in the study.
Figure 2. Flowchart of case ascertainment. Reprinted from study III, supplementary figure available online. PHAS = Public Health Agency of Sweden.
Inclusion criteria were: 1) that a sample drawn before symptom onset was available in either of the biobanks 2) that this sample was drawn before the age of 40 years, 3) that the individual later developed MS with a relapsing onset. For each such case-sample included, an additional sample from an individual that did not develop MS was selected as control-sample. To minimize confounding these controls were matched for biobank, sex, date of sampling and age at sampling.

Table 1. Characteristics of samples from cases and controls stratified by biobank.

<table>
<thead>
<tr>
<th>Biobank</th>
<th>Years of sample collection</th>
<th>Plasma or serum</th>
<th>N, (%) of a total 1340 samples</th>
<th>Proportion of samples from men/women (%)</th>
<th>Long-term storage temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Umeå</td>
<td>1976-2007</td>
<td>Both</td>
<td>204 (15)</td>
<td>21/79</td>
<td>-20°C</td>
</tr>
<tr>
<td>PHAS</td>
<td>1972-2001</td>
<td>Both</td>
<td>278 (21)</td>
<td>22/78</td>
<td>-20°C</td>
</tr>
<tr>
<td>Örebro</td>
<td>1994-2008</td>
<td>Serum</td>
<td>58 (4)</td>
<td>0/100</td>
<td>-20°C</td>
</tr>
<tr>
<td>Göteborg</td>
<td>1995-2009</td>
<td>Serum</td>
<td>94 (7)</td>
<td>11/89</td>
<td>-20°C</td>
</tr>
<tr>
<td>Skåne</td>
<td>1977-2007</td>
<td>Both</td>
<td>628 (47)</td>
<td>15/85</td>
<td>-20°C</td>
</tr>
<tr>
<td>Linköping</td>
<td>1993-2009</td>
<td>Both</td>
<td>78 (6)</td>
<td>10/90</td>
<td>-20°C</td>
</tr>
</tbody>
</table>

The biobanks represent a heterogenous material in terms of geographical catchment areas as well as demographics (Table 1). There were also differences, the extent of which are largely unknown, concerning handling and storage conditions of the blood samples. Most samples were serum (i.e. plasma without coagulations factors) but there were also some plasma samples in some of the biobanks. The time from sample draw until the sample was frozen varied, including both short-term storage in room temperature as well as cold storage. In the extreme cases, samples may have been refrigerated for up to a few months before being placed in freezers. Long-term storage temperature was the same for all biobanks at minus 20°C. A majority of samples (from all biobanks except PHAS) were intended for hepatitis or HIV screening (38%), followed by pregnancy screening (30%) and unspecified serological analysis (24%). PHAS conducts seroimmunity studies to evaluate the effects of vaccination programs, as well as directed surveillance during influenza pandemics. The samples generated from these activities are kept separate from the majority of samples contained in the biobank however, and it is likely that all samples that we obtained from PHAS originally were collected for the purpose of diagnostic microbiology (personal communication with PHAS representative).

Besides the pre-MS material used in all four studies, study I also included a much larger material consisting of individuals with established MS (n = 8742) and
matched controls (n = 7215), obtained from four different studies. Two included newly diagnosed individuals: Epidemiological Investigation of Multiple Sclerosis (EIMIS) and Stockholm Prospective Assessment of Multiple Sclerosis (SPASM/STOPMS), and two with prevalence design: Genes and Environment in Multiple Sclerosis (GEMS) and Immunomodulation and Multiple Sclerosis Epidemiology study (IMSE).

Biochemical methods

**Study I and IV**

A bead based multiplex assay\textsuperscript{105} was used to detect IgG antibodies against a selection of different viral proteins belonging to EBV, HHV-6A and HHV-6B. Importantly, to separate immune responses against the two HHV-6 viruses, a novel assay using especially species-divergent antigens was used. These were parts of the immediate early protein 1 (IE1) encoded by HHV-6A and HHV-6B (IE1A aa 382-638 and IE1B aa 400-775 respectively), as well as specific parts of the structural proteins p100 (aa 282-395) and 101k (aa 282-395). The HHV-6A encoded p100 antigen did not elicit enough of an immune response to be meaningfully differentiated from the background noise of the assay and was excluded from further analyses. Regarding antibodies against EBV, three different antigens were used. Two of these were parts of the protein EBNA-1, EBNA-1 trunc (aa 325-641) and EBNA-1 pep (aa 385-420). The third protein was the viral capsid antigen VCA p18 (aa 1-175). EBNA-1 pep is especially interesting in relation to MS as earlier studies have shown that individuals with MS have an especially high reactivity towards this EBNA-1 fragment compared to controls.\textsuperscript{106} These antigens were bacterially expressed as glutathione S-transferase fusion proteins in *E. coli* that in turn were lysed, after which the purified lysate was used to coat polystyrene beads.

The plasma and serum samples were diluted 1:1000 and pre-incubated to reduce background noise from unwanted antibodies against glutathione S-transferase or bacterial proteins. Then samples and mixed beads coated with different antigens were incubated in 96-well plates. In combination with the appropriate washing steps, biotinylated anti human IgG antibodies were added and detected by streptavidin–R–phycoerythrin (MOSS Inc., Pasadena, CA, USA), using a Luminex 200 analyser (Luminex Corp. Austin, Texas, USA). Reactivities against the different antigens were measured as median fluorescence intensity (MFI).
control samples did not react with the HHV-6A or HHV-6B antigens but the coefficient of variation (CV) for the EBV antigens were below 20%.

**Study II**
This study utilized two different commercial multiplex sandwich immunoassays (Meso Scale Discovery, Gaithersburg, MD 20877 USA) to analyse concentrations of leptin and insulin. One was a four-plex assay (K15174C) that, besides leptin and insulin, also included active GLP-1 and glucagon, and the other was a two-plex assay (K15164C) only including leptin and insulin. The four-plex assay was used in a pilot run on two 80-well plates, after which it was exchanged for the two-plex version, since neither active GLP-1 or glucagon was detectable in most samples and it required more sample volume.

Plates came pre-coated with capture antibodies from the manufacturer and a blocking solution containing a cocktail of proteins was applied in order to reduce non-specific binding and thereby reduce background noise. Combined with the appropriate washing steps, the samples and calibrators were incubated and after that, a detection antibody solution was added. Plates were analysed using a MESO QuickPlex SQ120 detecting electrochemiluminescence. The signal was then converted into concentrations of analytes using standards of known concentrations. Inter-plate CV was 5-6% for the two analytes.

**Study III**
This study applied liquid chromatography tandem mass spectrometry (LC-MS/MS) to measure concentrations of 25(OH)D₃ and 25(OH)D₂ (Figure 3). Briefly, serum samples were mixed with acetonitrile, containing 1% formic acid and internal standards (deuterated 25(OH)D₃ and D₂), to precipitate proteins. The mixtures were then passed through the columns of a Hybride SPE+ plate (Merck, Darmstadt, Germany) that both removes precipitated proteins and dissolved phospholipids. The extracts collected from the plates were evaporated to dryness, using nitrogen gas at 55°C, and were then reconstituted in a mixture of methanol/MilliQ-water (3:1). The samples were then analysed by LC-MS/MS using a Shimadzu Nexera LC system (Shimadzu Corporation, Kyoto, Japan) coupled to a Sciex QTrap 5500 MS (Sciex, Framingham, MA, USA). Concentrations of 25-hydroxyvitamin D₃ and 25-hydroxyvitamin D₂ were determined using calibration standards (seven levels between 0-350 nmol/L)
purchased from Chromsystems (Gräfelfing, Germany), and for quality assurance, internal controls (also from Chromsystems) were analysed with every batch of samples. Additionally, external controls from DEQAS (www.deqas.org) were analysed four times a year.

![Figure 3. LC-MS/MS analysis of a vitamin D standard containing 25(OH)D₃ and 25(OH)D₂. Combined chromatogram of specific mass transitions for 25(OH)D₃ and 25(OH)D₂ showing their internal separation as well as their separation from various minor peaks of inactive vitamin D epimers. CPS = counts per second.](image)

**Statistical methods**

All studies relied on logistic regression (conditional or unconditional) to calculate odds ratios (OR), with 95% confidence intervals (CI), as approximations of the relative risk to develop MS depending on the different exposures. For analyses where the dependent variable was continuous, such as antibody reactivity or analyte concentrations, linear regression was used instead. For comparisons of distributions between groups the Mann-Whitney U test or the Kruskal-Wallis test was applied, and correlations were investigated with Spearman’s Rho. Interactions were tested for by inclusion of interaction terms or by calculating the relative excess risk due to interaction (RERI) and attributable proportion (AP), as measures of departure from additivity of effects. Statistical calculations were performed in SPSS version 23 (IBM Corporation, New York, NY, USA) or the software R version 3.6.1 or version 3.4 (R Core Team, Vienna, Austria).
Statistical analyses were generally performed both in the entire material as well as in subgroups stratified by age at sampling or sex. Special consideration was given to age since a dominant hypothesis of MS etiology proposes that adolescence is an important age-period for determining an individual’s risk of developing MS. The age groups selected were <20 years, 20-29 years and 30-39 years. In most instances, an exception being study I, a case-control set was placed in the youngest or oldest group containing either a case or control, when cases and control were on different sides of the age cut-off. This was done to increase power in the groups with smaller numbers and enable the use of conditional logistic regression.

While study I and IV both investigated IgG antibodies against viral antigens, they had quite different approaches when it came to data analysis. In study I, a high and low immune response to the different HHV-6 antigens were used instead of an absolute cut-off for serostatus, since the assay used was not validated against a gold standard method and true serostatus thus could not be determined. High and low response was defined as being in the top and bottom quartiles among controls, and individuals with reactivity in the middle two quartiles were excluded from the logistic regression analysis. Study IV instead applied an arbitrary cut-off at 50 MFI to determine HHV-6A serostatus in an effort to maximise sensitivity while remaining above the limit of the technical noise of the assay at 30 MFI. EBV antigen serostatus was determined using previously published cut-offs validated against gold-standard reference assays, and analysed both separately and as a combined variable.

In study II and III more steps were taken to compensate for differences between biobanks as leptin, insulin and vitamin D may be less stable and/or more prone to be affected by pre-analytical handling and storage conditions than IgG. Leptin and insulin were transformed using a log base 10 transformation to achieve a more normal distribution (Figure 4) and thereafter converted to z-score, separately for men and women in each biobank. By converting concentrations to z-score, each data point is replaced by a value defined by how many standard deviations away from the mean it is located and provides a way to include men and women from all biobanks into one statistical model despite significant differences between the groups. In study III, 25(OH)D₃ concentrations were modelled as quintiles, with cut-offs defined among controls separately in each of the biobanks. This was done to compensate for differences between biobanks, where especially the PHAS biobank stood out with higher levels compared to the others, while still enabling inclusion of all data in the same statistical model.
Figure 4. Leptin and insulin concentrations after log base 10 transformation.

Ethical considerations

The regional medical ethical review board in Umeå approved of the original study (Dnr 2011-198-31M) and amendments (Dnr 2012-124-32M, Dnr 2013-226-32M, Dnr 2017-104-32M and Dnr 2018-468-32M). In addition, access to samples from the six biobanks was considered and approved by regional steering committees at each of the participating centres. Handling of sensitive information was approved by the Swedish Data Protection Authority, and the Swedish MS Society research board approved of the use of data from the Swedish MS Registry. Written informed consent was not required for participation, an exception being those individuals (~20%) that actively participated by donation of saliva for genotyping, as part of a related study. Potential study participants (n = 1365) were informed through a letter in the mail and given a chance to ask questions and opt out if they did not wish their already stored blood sample to be used. An address could not be found for a total of 92 individuals however, mainly due to their being deceased or having emigrated.
Using old biobank samples, some collected more than forty years ago, for research without written informed consent raises some important ethical issues. According to the Declaration of Helsinki (World Medical Association, 2013), informed consent must be sought before using such stored biological material, with the caveat that it might be permissible if such consent for some reason would be impractical or impossible to obtain. Further complicating matters, a majority of these samples were originally not intended for research purposes. While informed consent is required before collecting blood samples also in a clinical setting, it is doubtful that consent was given for long-term storage and use in future studies. On the other hand, these old samples represent a unique opportunity to do important research with relatively low risk for adverse effects among study participants. It is unlikely that individuals who already are afflicted with MS will benefit from this research but when also considering future potential patients it can be argued to be unethical not to use this available resource fully. It was in an attempt to navigate these difficulties and maximize scientific utility while also adhering to the ethical principles regulating medical research and protecting the integrity of participants, that an “opt-out” strategy was chosen. The legitimacy of this approach was supported by the low frequency of individuals declining use of their sample in this study (<1%).

The studies EIMS, IMSE, GEMS, and SPASM/STOPMS, from which the established MS material was supplied, were approved by the Regional Ethical Review Board in Stockholm and participants had provided written informed consent for inclusion in these studies.
Results

Characteristics of cases and controls

Study I was based on both pre-MS samples from three of the six biobanks, including a few unmatched cases and controls, as well as samples from a large number of individuals with established MS and matched controls. The other three studies were based around 670 complete case-control sets, although the exact number varied depending on sample availability for the different biochemical analyses. All pre-MS cases had a relapsing onset of the disease and 64% of cases with established MS had a relapsing disease course at the time of blood sampling. The median age at disease onset was similar between the two groups but the percent females were higher among those with samples taken before disease onset (Table 2).

Table 2. Demographics of cases and controls.

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-MS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of</td>
<td>676</td>
<td>674</td>
</tr>
<tr>
<td>% Female</td>
<td>84</td>
<td>84</td>
</tr>
<tr>
<td>Age [mean ± SD]</td>
<td>25.0 ± 6.4</td>
<td>25.0 ± 6.4</td>
</tr>
<tr>
<td>Median (IQR) age at MS onset</td>
<td>33 (12)</td>
<td>n.a.</td>
</tr>
<tr>
<td>Median (IQR) years from sampling until onset</td>
<td>8 (9)</td>
<td>n.a.</td>
</tr>
<tr>
<td>Established MS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of</td>
<td>8742</td>
<td>7215</td>
</tr>
<tr>
<td>% Female</td>
<td>72</td>
<td>75</td>
</tr>
<tr>
<td>Age [mean ± SD]</td>
<td>47.1 ± 14.0</td>
<td>48.2 ± 13.4</td>
</tr>
<tr>
<td>Median (IQR) age at MS onset</td>
<td>32 (15)</td>
<td>n.a.</td>
</tr>
<tr>
<td>Median (IQR) years from onset until sampling</td>
<td>11 (17)</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

One pre-MS case contributed two different samples and matched controls, one used in study I and the other in studies II-IV.

Study I - HHV-6A and HHV-6B

Samples from two different cohorts were analysed to determine IgG response against HHV-6A and HHV-6B antigens in this study. The larger cohort consisted of 8742 cases with established MS and 7215 matched controls and the smaller, pre-MS cohort, included 478 cases that later developed MS with relapsing onset and 476 controls. A total of 348 cases and 1 control that provided samples to the
established MS cohort were also included in the pre-MS material and were excluded from the larger cohort when similar statistical analyses were performed in both groups. Since unconditional logistic regression adjusted for age and sex was used, we opted to include a case-control set not matched for sex. This set was replaced with a correctly matched case-control set from another biobank, but with a later sampling date, in studies II-IV.

Being in the top quartile of antibody reactivity towards the HHV-6A encoded antigen IE1A was associated with MS (OR = 1.55; 95% CI: 1.42–1.69) as well as an increased risk of future MS (OR = 2.22; 95% CI: 1.54–3.19), compared to having a reactivity in the bottom quartile among controls (Table 3). This effect was most pronounced among young individuals, both in the established MS material (Age <30, OR = 1.80) and in the pre-MS group (Age <20, OR = 3.38). High reactivity toward the HHV-6B antigen IE1B was also associated with MS, but in the opposite direction (OR = 0.74; 95% CI: 0.67–0.81). There was no significant association with IE1B and case-status in the pre-MS group, however. As for the other HHV-6B antigen, 101K, there were significant associations with the risk of developing MS (OR = 1.51; 95 CI: 1.06–2.17) as well as sporadic associations with established MS, in some age-groups.

Additional analyses in the established MS material showed significant interactions between high antibody reactivity towards EBV and IE1A, as measured by departure from additivity (attributable proportion = 0.24; \( p = 6 \times 10^{-6} \)), while no such interaction was found between IE1A and Cytomegalovirus (CMV) antibody reactivity. There were also significant interactions between immunological response against IE1A and presence of the two major MS-risk HLA alleles, DRB1*15:01 and A*02:01 (AP = 0.31, \( p = 2 \times 10^{-8} \) and AP = 0.21, \( p = 2 \times 10^{-4} \) respectively. Associations between a high response against IE1A and IE1B and MS were similar among individuals with different disease courses, but a high response against 101K showed differential associations depending on if the case had RRMS or PPMS (OR = 1.18 and 0.66 respectively).

**Study II – Leptin and Insulin**

A total of 649 pre-MS case-control sets were included in this study. Median concentrations of leptin and insulin were similar among cases and controls in the entire material, as well as when stratified by age and sex. While insulin levels did
not significantly differ between the sexes, women had about five times higher median concentration of leptin compared to men (Figure 5.) There were no statistically significant associations between insulin levels and MS risk. A 1-unit increase in $z$-score leptin was associated with increased risk of developing MS among young individuals (<20 years) and among men (OR = 1.4; 95% CI: 1.1-1.9 and OR = 1.4; 95% CI: 1.0-2.0 respectively) (Table 3). For women, leptin was not significantly associated with MS risk however, but there were tendencies of lower risk with increasing age. While these associations of leptin and MS risk were not greatly affected by the inclusion of insulin in the logistic regression models, the association between increased leptin levels and decreased MS risk among women in the oldest age-group (30-39 years) became significant (OR = 0.74; 95% CI: 0.54-1.0).

![Figure 5. Median concentrations of insulin and leptin among cases and controls.](image)

**Study III – Vitamin D**

A total of 665 pre-MS case-control sets were analysed to determine concentrations of $25(OH)D_3$ and $25(OH)D_2$ in this study. Only 60 individuals had quantifiable levels of $25(OH)D_2$ (37 cases and 23 controls) and most statistical analyses were performed using $25(OH)D_3$. A majority were vitamin D replete with
54.6% of cases and 56.1% of controls having levels above 50 nmol/L. Seasonal variation was as expected with higher levels during the summer compared to the winter (Figure 6). Although there were no differences in median 25(OH)D₃ or total 25(OH)D between cases and controls, being in the top quintile of 25(OH)D₃ concentrations was associated with reduced risk of developing MS (OR = 0.68; 95% CI: 0.50–0.93). This finding was no longer significant in any of the age-groups (Table 3). There were no significant trends over 25(OH)D₃ quintiles, either in the entire material or when stratified by age (data not shown).

The PHAS biobank had the highest 25(OH)D₃ concentrations of all six biobanks, with a cut-off for the top quintile at 82 nmol/L. Sensitivity analyses excluding samples from PHAS and applying absolute cut-offs of 72 nmol/L (median for top quintile cut-offs among the other five biobanks) as well as using an even higher cut-off at 100 nmol/L yielded significant findings (OR = 0.69; 95% CI: 0.49–0.97 and OR = 0.30; 95% CI: 0.13–0.71 respectively). Analyses with 25, 50 and 75 nmol/L as cut-offs were not significant, however.

Figure 6. Median 25(OH)D₃ concentrations by month of sample collection. Data from 1330 samples with 48% (n = 633) collected in the winter months November-April.
Study IV - EBV

This study analysed IgG antibodies in samples from 670 pre-MS case-control sets, to determine reactivity against three different EBV antigens (EBNA-1 trunc, EBNA-1 pep and VCA p18) as well as the HHV-6A antigen IE1A and HHV-6B antigens IE1B and 101K. Cut-offs were applied to determine serostatus for each of the different antigens separately, but an individual was considered negative for EBV only if there was no seroresponse to any of the three EBV antigens. As there were no significant associations between either HHV-6B antigens and MS risk (Table 3), they were excluded from further analyses.

![EBV seropositivity](image)

**Figure 7.** EBV seropositivity among cases and controls.

A majority of cases and controls were seropositive against EBV and the proportion increased with increasing age. In the youngest age-group fewer cases than controls were positive (78% and 85% respectively) while the opposite was true for the older groups (Figure 7). Being seropositive against the EBV antigens EBNA-1 pep and VCA p18 and the HHV-6A antigen IE1A was associated with an increased risk of developing MS in the total cohort (OR = 1.6; 95% CI: 1.1-2.2, OR = 1.4; 95% CI: 1.0-2.0 and OR = 2.1; 95% CI: 1.6-2.7 respectively). Analyses stratified by age at sample collection revealed a pattern of differential associations in different age-groups for the EBV antigens, with a tendency for a lower risk among those below 20 years of age (Table 3). The interaction between EBV serostatus and age regarding MS risk was significant (p = 0.02). The age where seropositivity became a risk factor for MS was estimated at 18.8 years. There was
no significant additive interaction between EBV and IE1A seropositivity in relation to MS risk.

### Table 3. Overview of risk estimates for the studied MS risk factors

<table>
<thead>
<tr>
<th></th>
<th>All OR (95% CI)</th>
<th>&lt;20 years OR (95% CI)</th>
<th>20-29 years OR (95% CI)</th>
<th>30-39 years OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>aIE1A Q1 vs Q4</td>
<td>2.22 (1.54-3.19)</td>
<td>3.38 (1.46-7.81)</td>
<td>2.29 (1.43-3.67)</td>
<td>1.51 (0.65-3.49)</td>
</tr>
<tr>
<td>aIE1B Q1 vs Q4</td>
<td>0.96 (0.66-1.38)</td>
<td>1.09 (0.46-2.57)</td>
<td>0.88 (0.54-1.44)</td>
<td>1.09 (0.51-2.35)</td>
</tr>
<tr>
<td>a101K Q1 vs Q4</td>
<td>1.51 (1.06-2.17)</td>
<td>1.36 (0.58-3.15)</td>
<td>1.60 (1.00-2.54)</td>
<td>1.42 (0.64-3.14)</td>
</tr>
<tr>
<td>bLeptin</td>
<td>1.1 (0.97-1.3)</td>
<td>1.4 (1.1-1.9)</td>
<td>1.1 (0.94-1.3)</td>
<td>0.86 (0.67-1.1)</td>
</tr>
<tr>
<td>bInsulin</td>
<td>1.1 (0.94-1.2)</td>
<td>1.3 (0.96-1.7)</td>
<td>0.99 (0.85-1.1)</td>
<td>1.1 (0.84-1.4)</td>
</tr>
<tr>
<td>c25(OH)D3 Q1 vs Q5</td>
<td>0.68 (0.50-0.93)</td>
<td>0.60 (0.29-1.2)</td>
<td>0.70 (0.46-1.1)</td>
<td>0.72 (0.39-1.3)</td>
</tr>
<tr>
<td>dEBNA-1 trunc</td>
<td>1.3 (0.88-2.0)</td>
<td>0.59 (0.34-1.0)</td>
<td>2.9 (1.3-6.4)</td>
<td>6.5 (1.5-28.8)</td>
</tr>
<tr>
<td>dEBNA-1 pep</td>
<td>1.6 (1.1-2.2)</td>
<td>0.67 (0.39-1.1)</td>
<td>2.6 (1.5-4.7)</td>
<td>3.8 (1.6-9.4)</td>
</tr>
<tr>
<td>dVCA p18</td>
<td>1.4 (1.0-2.0)</td>
<td>0.91 (0.51-1.6)</td>
<td>1.7 (1.0-2.9)</td>
<td>2.1 (0.87-5.3)</td>
</tr>
<tr>
<td>dEBV</td>
<td>1.2 (0.72-1.8)</td>
<td>0.60 (0.32-1.1)</td>
<td>2.1 (0.87-5.3)</td>
<td>8.0 (1.0-64.0)</td>
</tr>
<tr>
<td>dIE1A 50MFI</td>
<td>2.1 (1.6-2.7)</td>
<td>2.0 (1.1-3.5)</td>
<td>2.0 (1.5-2.9)</td>
<td>2.5 (1.4-4.4)</td>
</tr>
<tr>
<td>dIE1B 50MFI</td>
<td>1.2 (0.93-1.5)</td>
<td>1.2 (0.76-2.1)</td>
<td>1.1 (0.81-1.6)</td>
<td>1.2 (0.76-1.9)</td>
</tr>
<tr>
<td>d101K 50MFI</td>
<td>1.2 (0.99-1.5)</td>
<td>1.3 (0.81-2.1)</td>
<td>1.3 (0.96-1.7)</td>
<td>1.1 (0.70-1.7)</td>
</tr>
</tbody>
</table>

- Unconditional logistic regression adjusted for age and sex, n = 954 (Study I).
- 1-unit increase in z-score value, n = 1298 (Study II).
- Quintiles defined separately for each biobank and then pooled, n = 1330 (Study III).
- n = 1340 (Study IV)

EBV positivity = seroresponse to either EBNA-1 trunc, EBNA-1 pep or VCA p18
MFI = Median Fluorescence Intensity, Q = Quantile, OR = Odds ratio, CI = Confidence Interval
Discussion

In summary, the main results of this thesis were the findings of significant associations between the risk of developing MS and HHV-6A, leptin, vitamin D and EBV. These findings lend further support to some of the already formulated theories of MS etiology, such as the hypothesized HHV-6A involvement in MS development and the increased MS risk that has been linked to adolescent obesity in earlier studies. They also indicate that this excess risk associated with obesity may be mediated by pro-inflammatory adipokines such as leptin independent of elevated insulin levels associated with metabolic dysfunction. Additionally, the already established protective effect on MS susceptibility from high serum concentrations of vitamin D was replicated in this study, but without evidence of a larger effect size among young subjects. Lastly, the old hypothesis that delayed EBV infection is a risk factor for MS, inspired mainly by the epidemiological similarities of MS with IM, is corroborated by the findings that EBV exhibits an age-dependency in relation to MS risk.

The results of this project were attained through the application of novel methods, such as a species-specific serology in study I and the use of metabolic markers in study II, and through the utilization of a unique study population that included a substantial number of young individuals. Both of these features, together with the fact that the case-samples were drawn before the onset of MS symptoms, were integral for the hypothesis testing that was the aim of this thesis. While most of the findings in this thesis support already formulated hypotheses there were also surprises, such as the inverse association between elevated leptin levels and MS risk among older women. A more thorough discussion of the results, merits and limitations of each specific study is provided in the following sections.

Study I – HHV-6A and HHV-6B

In this study we applied a novel bead-based assay, with the potential to differentiate between antibodies against proteins encoded by the two HHV-6 species, on a large MS and pre-MS case-control material. Our main finding was that individuals with MS, as well as individuals that later developed MS, had higher antibody responses towards the IE1 protein encoded by HHV-6A (IE1A), compared with healthy controls. Importantly, this effect was not seen for the IE1 protein encoded by HHV-6B (IE1B), which instead elicited less of a humoral
immune response among individuals with established MS compared with matched controls. Regarding IgG reactivities against the HHV-6B encoded structural protein 101K, the results were more ambivalent with some indications that immune responses towards this antigen differ depending on MS disease course.

These results strengthen the evidence of HHV-6A involvement in MS etiology, already suggested based on earlier studies using either PCR\textsuperscript{84,108–110} or immunofluorescence,\textsuperscript{111} capable of differentiating between the two HHV-6 species. While a serological approach, such as the one applied in this study, cannot be used to determine if active viral replication is taking place, it is reasonable to assume that an increased IgG reactivity may reflect an altered immune response due to increased viral load. This increased viral load could in turn be caused by differences between cases and controls regarding either the prevalence of infection, the intensity of primary infection or reactivation rates. Another possible explanation that must be considered is that genetic differences may be affecting how the immune system deals with these particular infections. This is unlikely to be the entire explanation, however. Addressing this issue, a GWAS was performed in the established MS cohort. Not surprisingly, significant associations of HLA variants with serological responses were found. When adjusting analyses for the most important HLA haplotypes, the earlier associations were mostly unaffected.

There are many potential mechanisms by which HHV-6A could be involved in MS development. One of the more general hypotheses of virus induced autoimmunity, centred around the idea that a virus could expose the immune system to host cell proteins by incorporating them into its envelope,\textsuperscript{112} may be worth exploring. In the specific case of HHV-6A and MS, this hypothesis would suggest that the virus triggers an immune response towards myelin components. This could happen if virions, with an envelope containing incorporated host cell proteins, were released from infected oligodendrocytes and subsequently processed by immune cells. Increasing the plausibility of this scenario are studies confirming the presence of HHV-6A in brain tissue\textsuperscript{113} as well as providing \textit{in vitro} support for ability of the virus to establish latent infection in oligodendrocytes.\textsuperscript{82} On the other hand it seems that HHV-6A has the capacity to reduce the capacity of dendritic cells to induce T-cell proliferation,\textsuperscript{114,115} a critical step in mounting an adaptive immune response.
Other ways for HHV-6A to partake in the chain of events leading up to MS development are through interaction with other viruses, such as EBV or endogenous retroviruses,\textsuperscript{116} or by directly affecting myelination. Regarding the latter option, there are indications that HHV-6A can interfere with the migration of oligodendrocyte progenitors\textsuperscript{117} which, while not being a cause of MS per se, would reduce the capacity for remyelination and thereby contribute to disease progression. Another possibility is that HHV-6A causes myelin disruption through directly cytotoxic effects on oligodendrocytes, a hypothesis supported by \textit{in vitro} data showing caspase independent cell death (i.e. necroptosis) induced by supernatants from infected cells.\textsuperscript{118}

The main limitation of this study is the fact that the assay used to determine serological responses against the different viral antigens could not be validated against a gold standard method. Because of this, we could not evaluate serostatus, which would have been valuable since much still is unknown about the epidemiology of HHV-6A. Additionally, the argument for primary HHV-6A infection as a driver of MS development, in large part, hinges on the ability to establishing different rates of infection among cases and controls. Another consequence of the lack of a gold standard method for validation is that we cannot rule out cross reactivity between the antigens derived from these two closely related viruses. An attempt to investigate this possibility was made by analysing samples from five children with primary HHV-6B infection, both during the acute phase as well as the convalescence phase. While none of the children had significant increases, (≥10 times MFI) in reactivity towards IE1A during convalescence, there were such responses against 101K and IE1B (80% and 40% respectively). This indicates that the assay indeed had the ability to detect an immune response against HHV-6B.

Despite aforementioned assay limitations, it is clear that cases and controls respond with different intensity towards these HHV-6A and HHV-6B derived antigens, and that these differences cannot easily be explained by genetic differences. If these findings are merely associations or if there are causal links between any of these viruses and MS development remains to be seen, however.
Study II – Leptin and Insulin

In this study we investigated the association between MS and obesity by measuring concentrations of two hormones linked to metabolic function, leptin and insulin, using an assay based on electrochemiluminescence. The main finding was an association between higher leptin levels and increased risk of developing MS among individuals below 20 years of age. Another interesting finding was the discrepancy between men and women, especially in regard to associations with MS risk in the different age-groups. Men showed associations of increased MS risk with increased leptin levels across all age-groups, that only became statistically significant when pooled. Women on the other hand displayed a tendency of lower risk of developing MS with increased leptin levels, as age increased.

The association between higher leptin levels and MS risk in the youngest age-group, seen when combining men and women in this study, is in line with earlier publications on obesity and MS, showing effects limited to childhood, adolescence and young adulthood.\textsuperscript{119} In contrast, there not only seem to be a lack of higher BMI among MS patients at disease onset,\textsuperscript{32} but actually some indications of a lower BMI, especially early in the disease course.\textsuperscript{27,120} BMI is a crude measure of adiposity however and a reason for lower BMI among individuals with MS could also be loss of lean tissue due to inactivity or the use of glucocorticoids. Interestingly, a recent cross-sectional study using dual-energy X-ray absorptiometry showed that fat mass among female MS patients followed a U-shaped pattern in relation to age, while total muscle mass constantly decreased with increasing age (Lautrup Frederiksen et al, ECTRIMS 2019, P1174). This indicates that loss of adipose tissue also could be a factor contributing to lower BMI among MS patients. Given that a substantial number older individuals in our study most likely already were affected by prodromal MS,\textsuperscript{121} it may also explain our surprising finding of a lower risk of MS with increased leptin levels among women aged 30-39, when adjusting for insulin. It does not however explain why this phenomenon was not seen among men and it is possible that leptin and/or obesity affects men and women differently. Why MS onset might induce fat loss is unclear, but it is not unreasonable to assume that brain inflammation affects systems responsible for energetic homeostasis. In line with this reasoning, there is evidence of weight loss from an experimental model of MS where the onset of EAE is preceded by an appetite suppressing surge in serum leptin.\textsuperscript{47}
Women have higher leptin concentrations than men and the present study was no exception, with about five times higher median levels of leptin among women compared to men. This sex difference seem to be influenced by at least two separate mechanisms, levels of fat mass and leptin secretion rate of adipose tissue, both of which are higher in women.\textsuperscript{122} If leptin indeed is a risk factor for developing MS, as the present study suggests, it is a risk factor that is present in much higher concentrations among females. It could thus prove to be an integral part in explaining why the disease is less common among men, as has proposed earlier.\textsuperscript{123} Leptin has also previously been put forward as a possible link between obesity and MS, and the fact that leptin functions as a mediator between adipose tissue and the immune system gives some credence to this notion.\textsuperscript{124} Especially incriminating as it relates to MS, are the effects of leptin on some of the immune cells currently considered to be fundamental in development of the disease, namely T and B lymphocytes. Interestingly, these effects are in some instances directly opposite those of active vitamin D (Figure 10).

There are important limitations relevant for this study that warrant discussion. Many relate to the non-standardized pre-analytical handling and potential effects of sample storage time on the blood samples. Especially insulin is sensitive to fasting status of the subject, where levels rise acutely after meals and we have no way of knowing if samples were drawn in a fasted state or not. Since controls were matched for biobank this applies equally for both cases and control, reducing the risk of systematic bias. Nevertheless, it may introduce random variations in insulin concentrations that make it more difficult to detect significant associations related to this biomarker. To estimate potential effects from the age of blood samples on leptin and insulin concentrations, we performed linear regression analyses. These showed that leptin levels decreased with sample age while there was no such effect on insulin concentrations (Figure 8). Thus, insulin seems not to be affected by long-term freeze storage. The increased leptin concentrations found in more recent samples was expected, given the rise in overweight and obesity during the last decades, but we cannot rule out an additional effect from degradation during long-term storage or from multiple freeze-thaw cycles. Another important limitation is the lack of data on BMI or other measurements that could have been used to approximate adiposity among the study subjects.
Despite these substantial limitations, we did find a significant effect of increased leptin levels on MS risk among young individuals, thereby confirming the findings of earlier studies estimating adiposity with BMI or body silhouettes. Interestingly, the same tendency was seen for increased insulin concentrations, although it did not reach statistical significance. This could be a result of the additional noise introduced into the data due to the lack of pre-analytical standardization (i.e. fasting status) or a lack of statistical power. Additionally, the fact that we did not have data on BMI or fat mass means that we cannot conclude if the effect seen for increases in leptin in this study is a direct effect of leptin, or if leptin simply is a marker for increased adiposity that in turn impacts MS risk through other mechanisms. It is also possible, although contradicted by the findings of Mendelian randomization studies, that obesity itself is a marker for something obesogenic in the environment that in itself directly influences MS risk.

**Study III – Vitamin D**

Before this study, there were only three studies utilising biobanked blood samples drawn before disease onset to investigate the connection between vitamin D and the risk of developing MS. Unique to this study however, was the relatively large number of young individuals (142 case-control sets below 20 years) and the use of LC-MS/MS to quantify 25(OH)D$_3$ and 25(OH)D$_2$. The main finding was that high levels of 25(OH)D, defined as the top quintile among controls, was associated with a reduced MS risk and the results are thus in line with those in the earlier studies.$^{94-96}$ We could however not reproduce the finding of a
significantly larger effect size among young individuals that has been previously reported.\textsuperscript{94}

The three earlier studies found significant effects of vitamin D on MS risk by using a wide variety of cut-offs, and the top quintile cut-off has also varied considerably between studies (41,\textsuperscript{96} 56,\textsuperscript{95} and 99,\textsuperscript{94} nmol/L respectively), depending on the absolute 25(OH)D levels among controls. We applied biobank specific cut-offs that varied in the comparatively high spectrum of 70-82 nmol/L. While it is tempting to try and compare absolute 25(OH)D levels between studies in an effort to tease out in what range there may be a protective effect, there are many reasons why such a comparison is problematic. One of the more straightforward reasons, that probably also contributes to the ongoing controversy of what constitutes optimal 25(OH)D levels, is differences between biochemical methods. A study comparing three different methods found that mean 25(OH)D concentrations differed depending on whether the method was mass spectrometry, radioimmunoassay or a chemiluminescent immunoassay (85, 70 and 60 nmol/L respectively).\textsuperscript{125} This translated to a difference of 8%, 22% or 43% of study subjects having insufficient 25(OH)D levels depending on method used, when applying a 50 nmol/L cut-off. Because methodologic differences such as these can have a large impact on results, it may be inappropriate to compare absolute 25(OH)D levels in our study that used gold standard method LC-MS/MS to the other three studies, using three different kinds of immunological assays.

We did not find large differences in 25(OH)D\textsubscript{3} between biobanks despite the apparent geographical differences. An exception was the PHAS biobank, collecting samples from the entire country, which had a greater proportion of individuals with high concentrations than the others. The reason for this discrepancy is unknown, but we speculate that the PHAS (the Swedish equivalent of the U.S. Centers for Disease Control and Prevention) is likely to receive an increased percentage of samples from individuals exposed to rare infectious diseases in countries closer to the equator. Such recruitment bias could make a pooled analysis applying absolute cut-offs problematic, and the findings of a 70% protective effect from having 25(OH)D\textsubscript{3} above 100 nmol/L was only significant in sensitivity analyses excluding this biobank.

Cutaneous production of vitamin D induced by solar UVB radiation is considered to be the main source of 25(OH)D globally, but dietary sources become relatively
more important farther away from the equator as the solar zenith angle, and with it UVB radiation, decreases with increased latitude (Figure 9). The intriguing latitude gradient of MS prevalence, implicating UVB/vitamin D in MS etiology, previously reported to be inversed in Scandinavia, has cast some doubts on the validity of this line of reasoning. In an updated analysis that included data from additional studies, this inversed gradient was no longer significant however, but the lack of such a gradient still requires an explanation. One hypothesis that has been put forth, and is supported by our results, is that serum levels of 25(OH)D are higher than would be expected from latitude alone in this geographical region, with for instance a large number of individuals in the northern part of Sweden being vitamin D replete, even in winter.

![Figure 9. Schematic image of vitamin D synthesis in relation to solar angle. For cutaneous production of vitamin D an angle of more than 45° is required, indicated by the shadow of an object being shorter than the object casting the shadow.](image)

There is an inverse correlation between serum 25(OH)D and BMI, likely due to dilution effects as 25(OH)D is fat soluble, which has inspired some to hypothesize that vitamin D is on the causal pathway of obesity in relation to the risk of developing MS. This is interesting considering that obesity not only decreases 25(OH)D but also results in increased leptin levels, which theoretically could have a, from an MS perspective, double negative effect on some of the immune cells believed to be important in MS etiology (Figure 10). Results from a recent Mendelian randomization study indicates that low vitamin D and high BMI both are independent risk factors for MS, however. This is in line with our findings that the protective effect of high vitamin D still was significant after adjusting for z-score leptin (data not shown).
**Figure 10.** Schematic representation of obesity in relation to vitamin D and leptin. The figure shows two ways in which obesity might influence the adaptive immune system to increase MS susceptibility. Small arrows (↑↓) indicate either positive or negative effects on proliferation/polarization or functional capabilities.\(^{44,88}\)

Important limitations of this study are that the included samples came from six different biobanks with distinct geographical catchment areas, different demographics and variations in other pre-analytical conditions. Fortunately, 25(OH)D\(_3\) has been shown to be stable for multiple freeze-thaw cycles as well as for refrigerated storage for up seven days. Exposure to daylight in room temperature may be more problematic however, as a decrease of 8.5% has been shown after seven days.\(^{134}\) A study of plasma stored at \(-20^\circ\text{C}\) showed that levels dropped from a mean of 73 to 63 nmol/l after four years,\(^{135}\) indicating that 25(OH)D is relatively stable also for long-term storage while simultaneously stressing the importance of matching for storage time when conducting studies using old samples. Another limitation is that only one sample per individual was available which means that the vitamin D concentrations measured are only a snapshot and may not be a reliable representation of levels over time.

By applying biobank-specific cut-offs of identical relative magnitude, this study gives further support for an involvement of 25(OH)D in MS etiology. It does not however, provide information on at what range of vitamin D concentrations a protective effect against MS may be located, but one can assume that it is likely
to be in the upper stratum of vitamin D distributions as indicated by the significant effect seen only for the top quintile in this study. Although the effect sizes converged with two of the earlier studies concerning a potentially stronger effect among young individuals, these results were not significant. This failure to replicate the finding of an increased effect among young individuals seen in one of the previous studies was surprising and may warrant further studies.

**Study IV – EBV**

In this study we aimed to test the hypothesis, put forward by Warner and Carp in 1981, that late in contrast to early infection with EBV is a risk factor for developing MS. To accomplish this, we applied the same multiplex assay as in study I, but to three separate EBV antigens. We also included additional data on antibody reactivity towards HHV-6A for comparison, as well as to investigate a potential interaction between these two viral infections in relation to MS risk. Our main finding was an apparent age dependency of EBV as a risk factor, with a significant trend of increased risk associated with seropositivity over the three age groups. No such trend was evident for HHV-6A and we found no significant interaction between the two viruses. Additionally, there was no indication of a larger effect size related to high IE1A reactivity in the youngest group compared to the older groups, in contrast to our findings in study I. This last finding was surprising, given the apparent effect of age seen previously, and was to a large part caused by the differing methodologies for determining high and low immune response in the two studies.

Infection with EBV is an environmental risk factor for MS with a large body of evidence associating it with the disease, and serological case-control studies have repeatedly shown that individuals with MS are more likely than controls to be seropositive for EBV.\(^{136}\) If only counting especially rigorous studies, defined as those applying two different methods to determine serostatus, 100% of MS patients seem to be positive.\(^{137}\) As for pre-symptomatic studies, there are some that have investigated this association as well.\(^{70,78,138–143}\) Results from these studies show that individuals that later develop MS have higher antibody reactivities against EBV antigens, especially EBNA, many years before symptom onset, indicating that these immunological responses not are a consequence of a dysregulated immune system due to the disease. While there were some indications of an age-dependency for EBV as an MS risk factor, none of these studies could really test the hypothesis that a late infection is a risk factor for MS.
while early infection is protective, mainly due to a lack of sufficient number of samples from young individuals.

While childhood EBV infections usually are asymptomatic, a primary infection after puberty often causes a powerful activation of the immune system with severe symptoms as a direct consequence. Studies linking a history of symptomatic EBV infection (i.e. infectious mononucleosis) to an increased MS risk provide indirect support for the importance of age at infection with EBV as it relates to MS and were instrumental in formulating this hypothesis in the first place. The differing associations between seropositivity to EBV and MS risk seen in the three age-groups in our study provide additional evidence for this presumed age-dependency of EBV infection. The tendency of a reduced MS risk associated with seropositivity before age 20, in combination with the reversed situation in older age-groups indicate that individuals that go on to develop MS generally are infected later in life than healthy controls.

An important limitation of this study is that we did not have multiple samples available from each individual and therefore could not determine the exact age of infection. This is especially problematic for the analyses of positivity among older individuals since we have no way of knowing at what age a case or control seroconverted. If a young individual is positive against EBV we can say with confidence that they were infected early in life, but it may be just as likely that an older individual who is positive also was infected at a young age. Fortunately, it is more straightforward to make assertions about those who were seronegative at a certain age, as they apparently not yet had developed a humoral immune response at that point in time. Another important limitation is that the age at sampling among cases was associated with the number of years between sampling and disease onset as well with the age at MS onset, which may have influenced the results.

While this study provides compelling serological evidence for an effect of age on EBV as a risk factor for MS development, no such effect was seen for HHV-6A. These results should be interpreted carefully however, since the assays ability to correctly discriminate between HHV-6A seropositive and seronegative individuals have not yet been established. Also, the lack of a significant additive interaction between the viruses may be due to power limitations, since some of the groups contained very few individuals.
General limitations

Although case-control studies are suitable to investigate the effects of exposures on rare outcomes, such as MS, they also have important limitations. They are a type of observational study and therefore cannot be used to determine cause and effect, only to show associations. Observational studies are also problematic since the findings may be a result of reverse causation, i.e. that the disease in itself causes the observed changes. In this study, the blood samples were collected before onset of MS and thus reducing, but not avoiding completely, this risk.

Another potential problem is selection bias. This term refers to a systematic difference between groups introduced during the selection process (of cases and controls), that is also associated with the exposure. An example of this would be if there was a systematic difference between cases and controls concerning the reason why they had samples stored in any of the biobanks included in this study, and that this differences was associated with one or more of the exposures studied. We cannot know to what extent this is the case, but one possibility is that a proportion of cases were affected by prodromal MS and that samples were drawn as part of a broader investigation of their symptoms. This would be in line with the recent findings of increased health care use in the years prior to a first demyelinating event.

Case-control studies are by nature retrospective since they begin with the process of identification of cases (i.e. instances of the outcome under investigation) and controls. This applies to this study as well, even though the samples that were used and therefore also the exposures measured, were collected prospectively. While all observational studies are subject to the risk of confounding (i.e. distortion of the true relationship between exposure and outcome due to a mixing of effects from additional factors) this risk can to some degree be addressed at the design stage. By matching cases and controls for variables that may contribute towards confounding (in this instance biobank, sex, date of sampling and age) we hope to have mitigated this risk. There are many possible confounders, both known and unknown, that have not been addressed, however.
Another question that warrants consideration is the extent to which the present project has external validity, that is how generalizable the findings made in this thesis are. It is possible that associations found in this sample of Swedish MS patients and healthy controls may not apply to other ethnic groups for instance. The presumed effect of vitamin D on MS susceptibility has been challenged recently based on, among other things, the lack of significant findings among blacks and Hispanics. There are also indications of the reverse, meaning that some MS risk factors such as smoking and IM may have a more pronounced effect in non-whites. Additionally, the cases included in the present study may not be representative even of Swedish MS patients, considering that there could be important differences between individuals that end up with samples stored in biobanks compared to those who do not.
Conclusions

I: Elevated antibodies against Human herpesvirus 6A, but not B, was consistently associated with increased multiple sclerosis risk.

II: Increased levels of leptin, but not insulin, was associated with increased multiple sclerosis risk.

III: High levels of vitamin D were associated with reduced risk of developing multiple sclerosis, but this effect was not more pronounced in young individuals.

IV: Epstein-Barr virus infection showed an age-dependency as a risk factor for multiple sclerosis. There was no interaction between infections with Epstein-Barr virus and Human herpesvirus 6A in relation to the risk of developing multiple sclerosis.
Future prospects

An important aspect in the investigation of risk factors for a disease such as MS that is presumed to be caused by an interplay of many different exposures, both environmental and genetic, is the concept of biological interaction. In a model of causality introduced by Rothman in 1976, he gives an explanation of how many different combinations of contributing factors may lead up the development of a specific disease, and that an individual’s risk of this disease can be described as the probability that a sufficient cause will emerge from component causes. In keeping with this model of causality it seems clear that none of the known risk factors for MS, including the ones studied in this thesis, by themselves are sufficient for MS development since they may be present in many individuals that never go on to develop the disease. It is likely however, that they act as separate components that together, in perhaps different combinations, can become sufficient to initiate disease. There is already some evidence of interactions between some of the known MS risk factors, which in turn indicates that these risk factors are part of the same causal pathway.

A logical next step is therefore to investigate potential interactions, both of the exposures included in this thesis, as well as results from additional analyses as they become available. One example of such an additional analysis, that already has been initiated, is the quantification of vitamin D binding protein. This protein has many diverse functions that may be of interest from a pathophysiological standpoint and such data would also enable estimations of the free fraction of vitamin D. We have also analysed inflammatory markers, including C-reactive protein which was associated with reduced MS risk among young individuals in a previous study, that could be used to adjust for systemic inflammation. Additional information from questionnaires regarding smoking, BMI and past sun exposure have also been collected and may, together with data on HLA risk alleles derived from saliva samples, help to further explore how environmental triggers can cause MS in susceptible individuals.
Acknowledgement

I would like to thank the following:

Main supervisor Peter Sundström for initiating the entire project and succeeding in the seemingly impossible task of convincing six different biobanks to part with a small volume of their valuable samples. Thanks also for being a role model and motivating me with your genuine enthusiasm for the field of MS research.

Co-supervisor Johan Hultdin for always taking the time to supervise my work, both in the lab and in front of the computer. Thanks also for being generous with your feedback and constantly reminding me to focus on what the data is telling us.

Co-supervisor Jonatan Salzer for being an inspiration and helping out whenever needed. Special thanks also for taking the time to sit through a much too long presentation and for providing feedback that helped improve it dramatically.

Jan Malm for taking an interest in the project and helping out with many different aspects all the way from start to finish. Thanks also for your generous feedback and encouragement.

Peter Andersen for providing valuable feedback and giving me some much-needed encouragement at critical moments.

Oluf Andersen, Daniel Jons, Lucia Alonso-Magdalena, Martin Gunnarsson and Magnus Vrethem for all your work identifying cases and for co-authoring the manuscripts.

Jörgen Andersson for technical support of all sorts. An especially big thanks also for going way beyond the call of duty in helping me prepare the send out to the study participants.

Kristin Nyman for keeping me company and reminding me to water the plants every week.

Andreas Tellström for showing me the ropes in the lab and working through weekends and holidays with me.

Staffan Lundstedt for performing the vitamin D analyses and for being so generous with your time when explaining the method.
Rob Buckland and Parviz Behnam for answering any questions I had about lab procedures and helping out with malfunctioning equipment.

Birgitta Nordin, Karin Moström, Agneta Valleskog, Nadia Ferry Yassin and all other administrators that keep the wheels turning at the university.

Göran Wadell, Stefan Söderberg and Torbjörn Lind for agreeing to be part of my midterm review group and providing some much-appreciated feedback.

Gabriel Granåsen for excellent statistical support.

Anders Nordström for running a pilot project and exploring alternative analytes and methods.

Elin Engdahl, Rasmus Gustafsson, Jesse Huang, Izaura Lima Bomfim, Pernilla Stridh, Mohsen Khademi, Nicole Brenner, Julia Butt, Angelika Michel, Maria Hortlund, Anna Karin Hedström, Louis Flamand, Masaru Ihira, Tetsushi Yoshikawa, Jan Hillert, Lars Alfredsson, Tim Waterboer, Tomas Olsson, Ingrid Kockum, Anna Fogdell-Hahn, Noemi Bender for all your hard work and for co-authoring one or more of the studies in this thesis.

Eva Ludvigsen and Esben Poulsen for helping out with everything related to the Mesoscale assays.

Special thanks also to everyone who donated blood to the biobanks and made this research project possible.
References


32. Hedström AK, Olsson T, Alfredsson L. High body mass index before age 20 is associated with increased risk for multiple sclerosis in both men and women. Mult Scler J 2012; 18: 1334–1336.


50. Petersen MC, Shulman GI. Mechanisms of insulin action and insulin


64. Dunmire SK, Hogquist KA, Balfour HH. Infectious Mononucleosis. In: Current topics in microbiology and immunology, pp. 211–240.


76. Derfuss T, Hohlfeld R, Meinl E. Intrathecal antibody (IgG) production
against human herpesvirus type 6 occurs in about 20% of multiple sclerosis patients and might be linked to a polyspecific B-cell response. *J Neurol* 2005; 252: 968–971.


139. Levin LI, Munger KL, Rubertone M V, et al. Temporal relationship between elevation of Epstein-barr virus antibody titers and initial onset of


