Title
Examining the Biological Mechanisms Underlying the Relationship Between Obesity and Multiple Sclerosis Susceptibility

Author
Gianfrancesco, Milena

Publication Date
2016

Peer reviewed|Thesis/dissertation

Examining the Biological Mechanisms Underlying the Relationship Between Obesity and Multiple Sclerosis Susceptibility

by

Milena Gianfrancesco

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Epidemiology

and the Designated Emphasis

in

Computational and Genomic Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Lisa F. Barcellos, Chair
Professor Alan E. Hubbard
Professor Nir Yosef

Summer 2016

Examining the Biological Mechanisms Underlying
the Relationship Between Obesity and Multiple Sclerosis Susceptibility

Abstract

Examining the Biological Mechanisms Underlying
the Relationship Between Obesity and Multiple Sclerosis Susceptibility

by

Milena Gianfrancesco

Doctor of Philosophy in Epidemiology

Designated Emphasis in Computational and Genomic Biology

University of California, Berkeley

Professor Lisa F. Barcellos, Chair

Background

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system that affects over 400,000 Americans and 2.5 million people worldwide. It is characterized by the presence of inflammation, neurodegeneration, and demyelinating lesions of white and gray matter. Both genetic and environmental factors have been implicated in MS etiology. Several genetic variants, including the human leukocyte antigen *HLA-DRB1*15:01* allele within the major histocompatibility complex (MHC) and 110 non-MHC variants, have been identified. Environmental risk factors associated with MS onset include exposure to tobacco smoke, Epstein-Barr virus or infectious mononucleosis, low levels of vitamin D, and most recently, obesity.

Obesity is a current public health problem around the world; approximately 35% of adults in the U.S. are obese. Further, obesity has more than doubled in children and quadrupled in adolescents over the past 30 years. Studies over the last decade have demonstrated that early childhood and adolescent obesity are significant risk factors for MS susceptibility. Therefore, the increasing prevalence of obesity could potentially be contributing to higher rates of MS in children and adults.

Methods

In this dissertation, I investigate the association between obesity and MS in several manners in order to understand the causal relationship and underlying biological relationship between these factors. In the first chapter, I examine whether self-reported body size during childhood and adolescence is associated with MS susceptibility, while controlling for several established genetic and environmental risk factors. In the second chapter, I utilize Mendelian randomization to estimate the causal relationship between obesity and MS using a weighted body mass index

genetic risk score (BMI GRS) of 97 variants previously identified to be associated with BMI. Additionally, I demonstrate evidence of protein-protein interactions between established gene regions associated with both BMI (n=97) and MS disease susceptibility (n=110), and that a subset of these are significantly associated with MS after adjusting for covariates. In the third chapter, I again examine the causal relationship between obesity and MS onset but instead focus on pediatric-onset MS (disease onset < 18 years). In addition, I examine whether a causal relationship between vitamin D and MS exists in pediatric-onset MS cases and controls, and whether BMI and vitamin D independently contribute to pediatric-onset MS susceptibility. Lastly, in chapter four, I identify genome-wide significant variants associated with serum levels of three adipokines to measure their relationship with MS: plasma soluble leptin receptor, adiponectin, and resistin.

Significance

This dissertation examines the complex contribution of obesity to MS susceptibility while accounting for other genetic and environmental risk factors for the first time. My findings establish strong evidence for potential underlying biological mechanisms between increased BMI and MS, identify potential genetic pathways that may be targeted for therapeutics, and indicate that interventions focusing on obesity prevention could in turn reduce the incidence of MS in the population.

Dedication

To my grandparents, whose sacrifices made it possible for this journey and inspired me every step of the way.

Table of Contents

Acknowledgements

First, I would like to acknowledge my advisor, Lisa Barcellos, for her support and guidance during my doctoral training over the past four years. Lisa's insightfulness and critical thinking have been fundamental to my development as a genetic epidemiologist, and I am grateful for her time and dedication to my education and future academic career.

Thank you to my other committee members, Alan Hubbard and Nir Yosef, for their valuable input on this dissertation. I am very grateful to Farren Briggs for his mentorship, advice, and inspiration. Thanks to collaborators at Kaiser Permanente Northern California, specifically Cathy Schaefer and Ling Shen, who provided the data used in this dissertation. Thanks to collaborators at UC San Francisco, including Maria Glymour, Emmanuelle Waubant, and Lindsey Criswell, for serving as great mentors and providing guidance for analyses included in this dissertation as well as additional projects. Special thanks to the GEGL research group for their insight, feedback and assistance relating to these chapters. I would also like to thank my fellow epidemiology classmates and the faculty at UC Berkeley and Yale School of Public Health, who have been supportive and encouraging, as well as Albert Lo and Elizabeth Triche for introducing me to the field of epidemiology.

Thank you to my grandparents, who encouraged the pursuit of education from an early age and acted as amazing examples of how hard work and dedication can lead to success. The numerous sacrifices they made throughout their lives made it possible for me to achieve all that I have and for that I am forever grateful. My grandmothers always served as strong female role models who pushed me to be the best I could, and never doubted my abilities.

Thanks to my family for always being there; even though you are three thousand miles away, you were my support system through it all and absolutely everything to me. Domenic, Daniella and Anthony- thank you for always keeping me grounded, laughing, entertained, and especially for being my best friends. Thanks to my father for showing me that the impossible is always possible, to never give up, and to always see things through to the end. Thanks to my mother for unconditional love and support, and reminding me that happiness is the most important part of life. I love you both so much.

Last, but not least, thank you to Arturo for being by my side through the ups and down over the past few years and for always believing in me, especially when I didn't believe in myself.

Portions of this dissertation work were supported by the NIH National Institute for Neurological Diseases and Stroke (NINDS) (F31 NS093832). I would like to acknowledge Elsevier for granting me permission to use their published material (License #3905970735446).

Chapter 1

Introduction

1.1 Motivation

Multiple sclerosis (MS) is a severe and complex disease of the central nervous system (CNS) that affects over 400,000 Americans and 2.5 million people worldwide.[1] Multiple sclerosis is characterized as an immune-mediated demyelinating disorder resulting in significant disability and decreased quality of life. It remains the second leading cause of neurological disability in young to middle-aged adults.[2] While advances in clinical management have been made over the past 30 years, long-term prognosis for most individuals diagnosed with MS remains poor.[3] After 20 years from onset, more than 60% of individuals with MS require ambulatory assistance; very progressive MS occurs in approximately 5-10% of individuals.[2,4]

Strong evidence supports the contribution of both genetic and environmental factors to disease susceptibility as demonstrated by increased, though incomplete, disease concordance among monozygotic twins (~25%) compared to dizygotic twins (~5%).[5-7] Substantial progress has been made towards the identification of several MS risk factors including the *HLA-DRB1*15:01* allele within the major histocompatibility complex (MHC).[8-11] The MHC region of the genome is compromised of a large cluster of genes found on chromosome 6 that play a major role in the immune system. However, studies have found that genetic risk of MS extends beyond this cluster. To date, 110 MS risk variants outside of the MHC have been recognized in individuals of European ancestry.[12]

Environmental risk factors have also been associated with MS onset. Three factors in particular have been consistently recognized in both cohort and case-control studies: exposure to tobacco smoke, Epstein-Barr virus (EBV) infection, and low levels of vitamin D.[13-15] The relationship between vitamin D and MS aligns with previous research demonstrating an association between MS and exposure to sunlight, as well as the geographic distribution of the 5 disease (low prevalence of MS along the equator). However, risk ratios for these factors are modest and the mechanisms underlying disease pathogenesis are still undefined. Research has indicated a role for gene-environment interactions, identifying susceptible subpopulations that may be at significantly greater risk. For example, smokers with two copies of a variant in the NAT1 gene are three times more likely to be diagnosed with MS compared to smokers without the variant, while nonsmokers are not affected by the variant.[16] Smoking has also been shown to interact with HLA genes to increase susceptibility to MS.[17]

Recently, obesity has emerged as a risk factor for MS. Association between body mass at age 18 or age 20 and MS onset later in life was observed in two studies,[18-19] where obese participants demonstrated greater than a twofold increase in risk of MS compared to those at normal weight. Additionally, childhood obesity and risk of both pediatric[20] and later onset[21] MS was reported. In pediatric cases, an increased risk was more prominent amongst females, and extremely obese children had over three times the odds of developing disease compared to normal weight children.[20] In a Danish study, childhood obesity was associated with 1.75 increased risk of

developing MS later in life when comparing BMI of girls > 95th percentile to girls < 85th percentile. [21] One study by Munger et. al (2009) did not find an association between obesity in childhood and risk of MS; however, this may be due to utilization of silhouette data to characterize body size during childhood, rather than body mass or weight calculation. Research has also indicated that HLA genes interact with BMI during adolescence to increase risk of MS.[22] All of these findings demonstrate that childhood, in addition to adolescence, is a particularly vulnerable period of exposure for MS risk, as has been suggested in the literature for obesity and other environmental factors such as sunlight exposure.[17,18,23]

The biological mechanism through which obesity and MS may be related is unknown; however, several hypotheses are plausible. Obesity is characterized by a chronic, low-grade inflammatory response supported by growing experimental evidence. Recent literature suggests that integration of metabolic tissue and immune cells contribute to obesity and obesity-related inflammation by sharing a common cellular target.[24] Alterations in adipose tissue in human studies may occur as early as in childhood.[25] Obesity during childhood and adolescence is also associated with increased levels of C-reactive protein, interleukin-6, and leptin levels,[26-27] indicating a proinflammatory state that may be important in MS pathogenesis. Interestingly, adverse serum lipid profiles have been associated with MS disease progression, and statins may be beneficial in early MS by reducing the migration of immune cells across the blood brain barrier.[28] The gut microbiota have also been reported to shape immune response and may influence peripheral inflammation.[29] One study found that gut bacteria influences neurologic inflammation through induction of Th-17 responses in experimental autoimmune encephalomyelitis, a well-established animal model for MS studies.[30] Whether individual gut microbiota contribute directly to inflammation, or instead act causally to influence the development of obesity, which in turn promotes inflammation, remains unknown.[31]

Additionally, adults and children with high body fat mass have lower circulating levels of vitamin D metabolites.[32-33] Lower levels of vitamin D have been associated with increased risk for MS[34] and more severe disease progression.[35] Therefore, overweight and obese individuals may be at particularly high risk for developing MS compared to normal weight individuals, especially during critical exposure periods of MS risk. Because vitamin D deficiency lies on the causal pathway between BMI and MS risk, studies are needed to determine whether high BMI or obesity confers risk of MS exclusively through vitamin D deficiency, or whether other mechanisms related to obesity are involved.

It is also plausible that there are common genetic and biologic pathways that contribute to obesity and result in susceptibility of MS. As noted above, MS is an inflammatory disease of the central nervous system. Similarly, new research has characterized obesity as an inflammatory disease; specifically, changes in immune cells and inflammation can result from obesity and high fat, high sugar diets.[23] Abnormal amounts of inflammatory protein PAR2 appears on abdominal fat tissues of obese humans and animal models of obesity. There may be a triggering or threshold effect of inflammation produced by both obesity and other factors that contribute to MS onset. In fact, some researchers hypothesize that individuals with genetic susceptibility to MS require the presence of some additional environmental factor to initiate disease.[36-37]

Approximately 100 genes associated with obesity-related traits have been cited in the literature through both animal and human studies. The most widely studied obesity gene is *FTO* (fat mass and obesity associated gene), with nine exons spanning more than 400 kilo-bases located on chromosome 16.[38] Many studies have shown that a cluster of variants in the first intron of *FTO*, which is highly conserved across species, are associated with obesity-related traits such as BMI, body weight, leptin levels, subcutaneous fatness, fat mass, as well as hip and waist circumference.[39] Studies have indicated that genetic variation in *FTO* at a particular single nucleotide polymorphism (SNP) rs9939609 can account for up to 6.6 pounds more in individuals with two copies of the minor allele (AA) compared to those with no copies (TT). Heterozygous carriers (AT) weigh on average 2.6 pounds more than individuals with zero copies (TT). Yet *FTO* only explains 1% of the population BMI variance, and 22% of the population attributable risk of obesity.[40] Beyond *FTO*, other obesity genes have been cited in the literature as being associated with obesity-related traits.[41] The mechanism of these genes and how they specifically contribute to obesity and disease processes, such as MS, remains to be investigated.

1.2 Specific Aims

This dissertation will examine the association between obesity and MS susceptibility using self-reported measures and genetic variants associated with BMI. I will utilize methods to estimate causality using observational data, and apply bioinformatics analyses to examine potential biological mechanisms underlying the relationship between obesity and MS. My specific aims are as follows:

1. To examine the relationship between self-reported BMI and MS onset (Chapter 2);
2. To estimate the causal relationship between increased BMI and MS onset using a genetic risk score composed of 97 variants (Chapter 3);
3. To explore the causal relationship between increased BMI and decreased vitamin D serum levels on pediatric-onset MS (Chapter 4); and
4. To investigate whether specific cytokines released by fat tissue are associated with MS onset (Chapter 5).

The first aim will use an MS case-control dataset from Kaiser Permanente, Northern California (KPNC). The second and fourth aims utilize the KPNC dataset in conjunction with an additional dataset of controls from the Kaiser Research Project on Genes, Environment and Health. My third aim uses a dataset comprised of pediatric-onset cases from across the United States and controls from multiple sources to increase statistical power of analyses. Findings from the second and third aim were additionally replicated in a dataset of MS cases and controls from the Karolinska Institute in Sweden. The final chapter concludes the dissertation by reviewing findings of the four studies conducted and proposes suggestions for future research.

Bibliography

1. Sadovnick AD, Baird PA, Ward RH. Multiple sclerosis: updated risks for relatives. Am J Med Genet. 1988;29(3):533-541.
2. World Health Organization. Atlas: Multiple Sclerosis Resources in the World. Multiple Sclerosis International Federation 2008.
3. Confavreux C, Vukusic S, Moreau T, et al. Relapses and Progression of Disability in Multiple Sclerosis. N Eng J Med. 2000;343:1430-1438.
4. Cottrell DA, Kremenchutzky M, Rice GP, et al. The natural history of multiple sclerosis: a geographically based study. 5. The clinical features and natural history of primary progressive multiple sclerosis. Brain. 1999 Apr;122(4):625-39.
5. Sadovnick AD, Baird PA. The familial nature of multiple sclerosis: age-corrected empiric recurrence risks for children and siblings of patients. Neurology. 1988;38(6):990-991.
6. Robertson NP, Clayton D, Fraser M, et al. Clinical concordance in sibling pairs with multiple sclerosis. Neurology. 1996;47(2):347-352.
7. Favorova OO, Kulakova OG, Boiko AN. Multiple sclerosis as a polygenic disease: an update. Genetika. 2010 Mar;46(3):302-13.
8. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med. 2007 Aug 30;357(9):851-862.
9. Patsopoulos NA, Esposito F, Reischl J, et al. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Annals of neurology. 2011 Dec;70(6):897-912.
10. Sawcer S, Hellenthal G, Pirinen M, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011 Aug 11;476(7359):214-219.
11. Barcellos LF, Sawcer S, Ramsay PP, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Hum Mol Genet. 2006 Sep 15;15(18):2813-24.
12. International Multiple Sclerosis Genetics Consortium (IMSGC). Analysis of immunerelated loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013 Sept <Epub ahead of print>
13. Ascherio A, Munger K. Epidemiology of multiple sclerosis: from risk factors to prevention. Semin Neurol. 2008 Feb;28(1):17-28.
14. Ebers GC. Environmental factors and multiple sclerosis. Lancet Neurology. 2008 Mar;7(3):268-277.
15. Lauer K. Environmental risk factors in multiple sclerosis. Expert review of neurotherapeutics. 2010 Mar;10(3):421-440.
16. Briggs FB, Acuna B, Shen L, et al. Smoking and Risk of Multiple Sclerosis: Evidence of Modification by NAT1 Variants. Epidemiology. 2014 <Epub ahead of print>
17. Hedstrom AK, Sundqvist E, Baarnhielm M, et al. Smoking and two HLA genes interact to increase the risk for multiple sclerosis. Brain. 2011;134:653-664.
18. Munger KL, Chitnis T, Ascherio A. Body size and risk of MS in two cohorts of US women. Neurology. 2009 Nov 10;73(19):1543-1550.
19. Hedstrom 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. 2012 Sep;18(9):1334-1336.
20. Langer-Gould A, Brara SM, Beaber BE, et al. Childhood obesity and risk of pediatric multiple sclerosis and clinically isolated syndrome. Neurology. 2013 Feb 5;80(6):548-552.

21. Munger KL, Bentzen J, Lauresen B, et al. Childhood body mass index and multiple sclerosis risk: a long-term cohort study. Mult Scler. 2013 Apr 2 [Epub ahead of print].
22. Hedstrom AK, Bomfim L, Barcellos L, et al. Interaction between adolescent obesity and HLA risk genes in the etiology of multiple sclerosis. Neurology. 2014;82(10):865-872.
23. Islam T, Gauderman WJ, Cozen W, et al. Childhood sun exposure influences risk of multiple sclerosis in monozygotic twins. Neurology. 2007 Jul 24;69(4):381-8.
24. Lim J, Iyer A, Liu L, et al. Diet-induced obesity, adipose inflammation, and metabolic dysfunction correlating with PAR2 expression are attenuated by PAR2 antagonism. FASEB. 2013;27:4757-4767.
25. Sbarbati A, Osculati F, Silvagni D, et al. Obesity and inflammation: evidence for an elementary lesion. Pediatrics. 2006 Jan;117(1):220-223.
26. Chu NF, Chang JB, Shieh SM. Plasma leptin, fatty acids, and tumor necrosis factorreceptor and insulin resistance in children. Obes Res. 2003 Apr;11(4):532-540.
27. Kanneganti TD, Dixit VD. Immunological complications of obesity. Nat Immunol. 2012 Aug;13(8):707-712.
28. Weinstock-Guttman B, Zivadinov R, Mahfooz N, et al. Serum lipid profiles are associated with disability and MRI outcomes in multiple sclerosis. J Inflamm. 2011; 8:127.
29. Kau AL, Ahern PP, Griffin NW, et al. Human nutrition, the gut microbiome and the immune system. Nature. 2011 Jun 16;474(7351):327-336.
30. Lee YK, Menezes JS, Umesaki Y, et al. Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis. Proc Natl Acad Sci USA. 2011 Mar 15;108 Suppl 1:4615-4622.
31. Tsai F, Coyle WJ. The Microbiome and Obesity: Is Obesity Linked to Our Gut Flora? Current Gastroenterology Reports. 2009;11:307-313.
32. Parikh SJ, Edelman M, Uwaifo GI, et al. The relationship between obesity and serum 1,25-dihydroxy vitamin D concentrations in healthy adults. J Clin Endocrinol Metab. 2004 Mar;89(3):1196-1199.
33. Smotkin-Tangorra M, Purushothaman R, Gupta A, et al. Prevalence of vitamin D insufficiency in obese children and adolescents. J Pediatr Endocrinol Metab. 2007 Jul;20(7):817-823.
34. Munger KL, Levin LI, Hollis BW, et al. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. JAMA. 2006 Dec 20;296(23):2832-2838.
35. Mowry EM, Waubant E, McCulloch CE, et al. Vitamin D status predicts new brain magnetic resonance imaging activity in multiple sclerosis. Ann Neurol. 2012 Aug;72(2):234-240.
36. Goodin, D.S., The causal cascade to multiple sclerosis: a model for MS pathogenesis. PloS One. 2009;4(2):e4565.
37. Handel, A.E., et al., Environmental factors and their timing in adult-onset multiple sclerosis. Nature reviews. Neurology. 2010;6(3):156-66.
38. Scuteri A, Sanna S, Chen W, Uda M, Albai G, Strait J, et al. Genome-Wide Association Scan Shows Genetic Variants in the FTO Gene Are Associated with Obesity-Related Traits. PLoS Genet 3. 2007; 3(7):1200-1210
39. Loos RJF, Bouchard C. FTO: the first gene contributing to common forms of human obesity. Obesity Reviews. 2008;9(3):246-250.

40. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nature Genetics. 2007;39(6):724-726.
41. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015 Feb 12;518(7538):197-206.

Chapter 2

Obesity during childhood and adolescence increases susceptibility to multiple sclerosis after accounting for established genetic and environmental risk factors

2.1 Background

Multiple sclerosis (MS) is a severe and complex disease of the central nervous system (CNS) that affects over 400,000 Americans and 2.5 million people worldwide.[1] It remains the second leading cause of neurological disability in young to middle-aged adults.[2] While advances in clinical management have been made over the past 30 years, long-term prognosis for most individuals diagnosed with MS remains poor.[3] After 20 years from onset, more than 60% of individuals with MS require ambulatory assistance; very progressive MS occurs in approximately 5-10% of individuals.[2,4] Strong evidence supports the contribution of both genetic and environmental factors to disease susceptibility as demonstrated by increased, though incomplete, disease concordance among monozygotic twins (~25%) compared to dizygotic twins (~5%).[5-7] Substantial progress has been made towards the identification of several MS risk factors including the *HLA-DRB1*15:01* allele within the major histocompatibility complex (MHC) and other non-HLA genetic variants,[8-10] as well as exposure to tobacco smoke, Epstein-Barr Virus (EBV) infection, and lower levels of vitamin D;[11-13] however, the mechanisms underlying disease pathogenesis are still undefined.

Recently, obesity has emerged as an important risk factor for MS. Association between body mass at age 18[14] or age 20[15] and MS onset later in life was observed in two studies, where obese participants demonstrated greater than a twofold increase in risk of MS compared to those at normal weight. Additionally, childhood obesity and risk of both pediatric[16] and later onset[17] MS was reported. In pediatric cases, an increased risk was more prominent amongst females, and extremely obese children had over three times the odds of developing disease compared to normal weight children.[16] In a Danish study, childhood obesity was associated with 1.75 increased risk of developing MS later in life when comparing BMI of girls $\geq 95^{th}$ percentile to girls $< 85^{th}$ percentile.[17] While obesity has been shown convincingly to be associated with MS, previous studies have not investigated this relationship while accounting for effects of other established risk factors. As both childhood and adolescence may be critical exposure periods for MS,[18] this study aimed to examine whether body size/mass during childhood and adolescence were associated with MS while controlling for a number of environmental and genetic risk factors, including history of infectious mononucleosis and *HLA-DRB1*15:01* status, the strongest genetic contributor to MS.[19]

2.2 Methods

Participants
Data were collected from members of Kaiser Permanente Medical Care Plan, Northern California Region (KPNC). Both incident and prevalent MS cases were studied. KPNC is an integrated health services delivery system with a membership of 3.2 million that comprises about

25-30% of the population of a 22 county service area and is the largest healthcare provider in northern California. Comparisons with the general population have shown that the membership is objectively representative; however, persons in impoverished neighborhoods are underrepresented.[20] The membership is stable with 64% of all members, and over 72% of those aged 40 or more years, maintaining membership for five years or more. Individuals with chronic conditions such as MS have historically been more likely to remain members. The KPNC MS Research Program was recently established to support epidemiologic investigations of both genetic and environmental risk factors in a large, population-based study sample.

Eligible KPNC cases were defined as: individuals with a diagnosis of MS by a neurologist (*multiple sclerosis,* ICD9 code 340.xx; 95% had at least two MS diagnoses by a neurologist), current age of 18 through 69 years, and membership in KPNC at initial contact. The study was restricted to white (non-Hispanic) race/ethnicity, the population with the highest prevalence of MS. The treating neurologist of each MS case was contacted for approval to contact each case as a potential MS study participant. Eligibility for the study was confirmed at initial contact, and diagnoses were validated utilizing medical record review and published diagnostic criteria.[21-22]

Controls were KPNC current members without a diagnosis of MS or related condition (optic neuritis, transverse myelitis, or demyelination disease; ICD9 codes: 340, 341.0, 341.1, 341.2, 341.20, 341.21, 341.22, 341.8, 341.9, 377.3, 377.30, 377.39, and 328.82) confirmed through electronic records, and white (non-Hispanic) race/ethnicity. Potential study participants were contacted by mail with a follow-up phone call to explain the study and procedures. The average participation rate was 58% for controls and 79% for cases. A total of 1,932 individuals (1,235 MS cases and 697 controls) with data on body size were studied at the time of the data freeze (February 2013). Study protocols were approved by the Institutional Review Boards (IRB) of KPNC and the University of California, Berkeley.

KPNC Exposure Assessment

KPNC participants completed a computer-assisted telephone interview (CATI) administered by trained staff interviewers and comprised of questions related to various events and exposures. Education level was defined as the self-reported highest education level attained on an 8-point scale: “none”; “grade school only (1-8)”; “some high school”; “not high school graduate”; “high school graduate or GED”; “some college or technical/trade/vocation school or associate’s degree”; “bachelor’s degree”; “master’s degree”; and “doctoral degree.” Smoking was classified as ever or never based on: “Have you ever smoked at least one cigarette per day for one month or more?” Sun exposure at 10 years of age was assessed by asking “At 10 years of age, how often did you sunbathe in the summer (lay in the sun with a bathing suit on between the hours of 10am-2pm)?” and categorized as: ““almost every day”; “2-5 days per week”; “at least once per week”; “1-2 times per month”; and “never”. Physical activity at 10 years of age was determined by inquiring: “Overall, as a young girl/boy at 10 years of age you were:” with possible responses being “not physically active”; “a little physically active”; “moderately physically active”; and “very physically active”. Additional information included residence at birth and age 10 (city, state), birth weight, having been breastfed as an infant, mother and father’s body size at age 30, and history of infectious mononucleosis as a proxy for pre-MS EBV serostatus, which was not available for study participants. EBV exposure can manifest as infectious mononucleosis[23] and

previous epidemiologic studies have repeatedly shown association between this condition and the development of MS.[24] Onset year of MS, determined as year of first self-reported symptom, and age of diagnosis were determined based on a series of standardized CATI questions including: "How old were you when you had your first symptoms of MS?", "How old were you when a doctor first told you that you had MS?" Year for symptom onset and age of diagnosis were calculated using date of birth provided in the electronic medical record (EMR). Additional questions such as "Prior to being diagnosed, which of the following describes your very first MS symptoms which lasted for 3 or more days?" (13 possible first symptoms were defined). When possible, EMR data were used to validate self-reported symptoms.

KPNC study participants reported their current weight and height at the time of the interview, as well as their highest and lowest (non-pregnancy) weight during their 20's, and their 30's. Self-reported height and weight have repeatedly been shown to be valid for identifying relationships in epidemiologic studies, as self-reported and measured weights show strong correlation.[25-26] Previous studies have also demonstrated that recalled weight at 18 years of age and self-reported height are highly valid,[27] including Troy et al. (1995) in which women aged 25-42 in the Nurses' Health Study II were examined.[28]

Body mass index (BMI) for each participant was calculated by dividing weight in kilograms by height reported at time-of-interview in meters squared. Mean BMI of each participant was calculated by averaging the highest and lowest BMI during each period. BMI categories were divided according to the World Health Organization's definitions: <18.5 kg/m^2 (underweight), 18.5-<25 kg/m^2 (normal weight), 25-<30 kg/m^2 (overweight), and $\geq$ 30 kg/m^2 (obese). The categories for normal and overweight were subdivided to measure smaller variations in MS risk analogous to previous studies: <18.5, 18.5-<21, 21-<23, 23-<25, 25-<27, 27-<30, and $\geq$ 30.[14-15]

KPNC participants were asked to recall body size at age 10 and age 20 from one of four categories ("underweight," "just about right," "little overweight," or "very overweight"), with the two overweight categories combined for analyses. Participants also identified their body type at time-of-interview, age 10, age 20, and age 30 from one of nine body silhouettes, which ranged from very thin to extremely obese.[29] The largest four body type categories were combined for consistency with prior studies[14] and to avoid small sample size in these categories. At the time of the data freeze, silhouette body type information was obtained for 72.5% of the dataset (906 cases and 496 controls). Missing data were included as an indicator variable.

Whole blood was collected, processed and extracted for DNA using Gentra Puregene protocol. Saliva was collected using Oragene kits. Medium resolution *HLA-DRB1* and genome-wide SNP genotyping was performed as previously described.[10,19,30] Additionally, a weighted genetic risk score (wGRS) was calculated for each individual that combines the weighted OR from each of 110 non-MHC MS susceptibility loci identified through recent GWAS and follow-up studies.[10,31,32] The wGRS was calculated by multiplying the number of risk alleles for each locus by the weight for that variant and then taking the sum across the 110 loci. Genetic data were available for 88% of study participants (86% of males; 89% of females).

Statistical Analyses

Demographic differences between cases and controls were compared using X^2 test and

independent sample t-test where appropriate; Fisher's exact test was utilized in instances with small cell counts. Stratified analyses were performed for female and male cases and controls due to conflicting evidence regarding the relationship between BMI and MS by gender.[15-17] Logistic regression models were used to study multiple risk factors in addition to BMI in KPNC cases and controls. The primary predictor of each multivariate model was BMI/ body size at various ages, adjusted for year of birth, self-reported history of cigarette smoking, and college education. Tests for linear trend across BMI categories were assessed by modeling BMI categories as continuous, ordinal variables. Both crude and adjusted ORs with 95% CIs were estimated.

When individually added to the multivariate model, latitude at birth and age 10 (calculated from city and state of residence at these time points), physical activity at age 10, sun exposure at age 10, having been breastfed as an infant, birth weight, mother's body size at 30, and father's body size at 30 did not significantly contribute to the models (P>0.10) and thus were not included in subsequent analyses. Self-reported history of infectious mononucleosis, number of copies of (*HLA)-DRB1*15:01*, as well as individual wGRS, were included in final models to investigate if body size/mass contributed to MS risk after adjusting for these established risk factors.

To minimize the potential for reverse causality of MS on body mass and body size, analyses were restricted to MS cases where age of first symptom was > 15 years old for variables involving childhood (39 MS cases excluded; N=1,196 cases), ≥ 30 years old for variables involving BMI during 20's (an additional 499 cases excluded; N= 697 cases), and ≥ 40 for variables involving BMI during 30's (an additional 431 cases excluded; N= 266 cases). Spearman's rank correlation coefficients were computed to measure the relationship between body size variables at various ages, and Wald tests were conducted to measure differences between full and restricted models. All analyses were conducted in Stata v11.2 (StataCorp, TX). This study was focused on a single hypothesis established a priori. A type I error of 5% (α=0.05) was utilized for significance. Attributable risk percent for measures of body size was calculated as: [(OR-1/OR)*100].

2.3 Results

Demographic differences between female MS cases and controls (986 cases, 585 controls) were found with respect to year of birth and education (Table 1). Further, smoking status, ever having infectious mononucleosis, *DRB1*15:01* status and wGRS were strongly associated with MS. There was no significant difference in BMI at time-of-interview in females. Within males, there were significant differences between cases and controls (249 cases, 112 controls) for BMI at time-of-interview, *DRB1*15:01* status and wGRS.

Female report of being a little/very overweight during childhood was weakly correlated with report of being a little/very overweight in one's twenties (Spearman's rank correlation coefficient ρ = 0.34); mean BMI in one's 20's (ρ = 0.23), and mean BMI in one's 30's (ρ = 0.23). Body silhouette report at age 20 and at age 30 were also weakly correlated with body silhouette report at childhood (ages 10 & 20 ρ = 0.35; 10 & 30 ρ = 0.33; 20 & 30 ρ = 0.33). Wald tests determined that statistical models including body size variables at all ages were not significantly different from models including only a single body size variable (P>0.10), thus results were reported only for restricted models.

Being a little/very overweight at age 10 was significantly associated with MS (P =3.00 x 10^{-3}), as was report of being a little/very overweight at age 20 in females (P=2.50 x 10^{-3}) (Table 2). Estimates trended in the same direction for males, but were not significant. For females, increased odds of MS were significantly associated with BMI in 20's for categories 21-<23 kg/m^2 (1.39, 95% CI 1.02, 1.91), 25-<27 kg/m^2 (1.77, 95% CI 1.06, 2.97), and ≥ 30 kg/m^2 (2.15, 95% CI 1.18, 3.92) (P-value for trend = 9.60 x 10^{-4}). No associations between BMI in 30's and MS were observed in females within any BMI category. Significant associations between BMI in 20's and MS in males were not observed; however, some evidence of a protective effect for BMI in 30's was revealed (P-value for trend = 0.04).

Multivariate models examining body size at age 10 and susceptibility to MS in females, and restricted to individuals with complete genetic information, demonstrated a consistent and more pronounced association after controlling for conventional covariates (age, smoking, and education), history of infectious mononucleosis, and various genetic factors relating to MS predisposition (Table 3a). Similarly, this association was evident for body size at age 20 (Table 3b) and mean BMI in 20's (Table 3c) in females. There was no association between body size/BMI at any age period and MS in males when restricting to individuals with complete genetic information and controlling for conventional covariates. Associations remained insignificant when infectious mononucleosis and genetic risk factors were added to the model (data not shown).

When variables of body size during childhood and mean BMI in twenties were considered in the same model and multivariate analyses restricted to female MS cases with age at onset between 20 and 30 years old, report of being a little/very overweight at age 10 was significantly different from controls (P=0.01), while being little/very overweight at age 20 was not (P=0.68) (Figure 1). In contrast, female MS cases with age at onset 30 years old or later, being a little/very overweight at age 10 was not significant (P=0.14), and report of being a little/very overweight at age 20 was significant (P=0.049).

Body size as reported by silhouette identification was not significantly different between female cases and controls at time-of-interview (Supplementary Table 1). Report of larger body size at age 10, however, was associated with increased odds of MS (adjusted OR body size 5 vs. 3 = 1.59, 95% CI 1.01, 2.49); report of larger body size at age 20 also trended toward significance (adjusted OR body size 5 vs. 3 = 1.58, 95% CI 0.99, 2.52). No association between body size as reported by silhouette identification was found in KPNC males at any age (data not shown).

2.4 Discussion

The current study is the very first to demonstrate a significant association between self-report of being overweight at age 10, as well as at age 20, and MS in females after controlling for established genetic and environmental risk factors. Findings demonstrated a significant and two-fold increased risk of MS with increasing BMI similar to previous findings of prospective cohorts.[14] Additionally, the odds of MS in females increases linearly with BMI, providing strong evidence to support a dose effect. The observed association between body mass and MS persisted after controlling for history of infectious mononucleosis and smoking, *HLA-*

*DRB1*15:01* status, and the combined effects of known non-MHC risk alleles in females. Notably, the magnitude of BMI association with MS is strong and similar to other identified genetic and environmental MS risk factors. We also compared our findings to those from the Epidemiological Investigation of MS (EIMS), a large MS population-based case-control study in Sweden.[15,33] Association between BMI at age 20 and MS was examined in 1,571 MS cases and 3,371 controls matched on age, gender and residential area. Similar to the current study, an increased risk of MS was associated with increasing BMI at age 20 in females (Supplementary Figure 1), providing additional support for the importance of this time period in MS susceptibility.

The growing obesity epidemic significantly impacts public health at both local and global levels. Current estimates show more than one third of adults and approximately 17% of children in the United States are obese.[34] Common disorders such as cardiovascular and metabolic diseases, as well as many cancers have been convincingly linked to obesity.[35] Obesity also has been recently established as a risk factor for a number of chronic and autoimmune diseases, including MS.[36] Results indicate that given a causal relationship, approximately 33% of MS cases can be attributed to being a little/very overweight at age 10, amongst females reporting this weight status.

Our findings demonstrate that childhood, in addition to adolescence, is a particularly vulnerable period of exposure for MS risk, as has been suggested in the literature for obesity and other environmental factors such as sunlight exposure.[14,15,18] Results suggest that body size during the period *immediately preceding onset of symptoms* may be an important factor for MS susceptibility in females; however, further investigation is warranted.

Investigations of childhood obesity and risk of MS have reported conflicting findings. A recent study found that extremely obese children had over three times the odds of developing pediatric MS compared to normal weight children, with risk especially strong amongst females.[16] A previous study by Munger et. al (2009) did not find an association between obesity in childhood and risk of MS, though this may be due to utilization of silhouette data to characterize body size during childhood. While strong associations between MS and being a little/very overweight during childhood and adolescence were observed in the current study, results based on silhouette data were not consistent. This may be due to a more favorable perception of body silhouettes in overweight individuals,[37] which would bias results towards the null, and reduced available power to detect an association.

While childhood and adolescent body size in males suggested an increased risk of MS, results were not significant. Previous studies examining this relationship in males have not shown an association between obesity and pediatric MS,[16] or between childhood obesity and MS with a later onset;[17] however, risk of MS amongst males was greater with increasing BMI at age 20 in the EIMS study.[15] In contrast, our results demonstrated a null association between MS and BMI in 20's, and an inverse relationship between MS and increasing BMI in 30's. When analyses were restricted to males with genotype data, the inverse association was not significant and persisted after controlling for history of infectious mononucleosis and established genetic risk factors. Thus, further investigation of BMI and MS in males is needed. It should be noted that our sample size was small and wide confidence intervals were observed. Larger studies are

needed to determine whether obesity is a strong independent risk factor for males, as demonstrated for females.

The biological mechanism through which obesity and MS may be related is unknown; however, several hypotheses are plausible. Obesity is characterized by a chronic, low-grade inflammatory response supported by growing experimental evidence. Recent literature suggests that integration of metabolic tissue and immune cells contribute to obesity and obesity-related inflammation by sharing a common cellular target.[38] Alterations in adipose tissue in human studies may occur as early as in childhood.[39] Obesity during childhood and adolescence is also associated with increased levels of C-reactive protein, interleukin-6, and leptin levels,[40,41] indicating a proinflammatory state that may be important in MS pathogenesis. Interestingly, adverse serum lipid profiles have been associated with MS disease progression, and statins may be beneficial in early MS by reducing the migration of immune cells across the blood brain barrier.[42,43] The gut microbiota have also been reported to shape immune response and may influence peripheral inflammation.[44] One study found that gut bacteria influences neurologic inflammation through induction of Th-17 responses in experimental autoimmune encephalomyelitis, a well-established animal model for MS studies.[45] Whether individual gut microbiota contribute directly to inflammation, or instead act causally to influence the development of obesity which in turn promotes inflammation, remains unknown.[46] A recent study also demonstrated an association between self-reported abuse during childhood and risk of severe obesity later in life.[47] Future studies examining obesity as a mediator of stressful life events and MS may be informative.

Additionally, adults and children with high body fat mass have lower circulating levels of vitamin D metabolites.[48,49] Lower levels of vitamin D have been associated with increased risk for MS[50] and more severe disease progression.[51] Therefore, overweight and obese individuals may be at particularly high risk for developing MS compared to normal weight individuals, especially during critical exposure periods of MS risk. The current study could not assess individual pre-disease levels of serum vitamin D; however, we do not see this as a limitation, similar to previous studies.[14] While information on time spent in the sun at age 10 was included as a potential proxy for sun exposure and resulting vitamin D levels, no association with MS was observed. One interview question was used in the current study to assess sun behavior at age 10. A more extensive index of sun exposure at this important time period might be more informative. Because vitamin D deficiency lies on the causal pathway between BMI and MS risk, further studies are needed to determine whether high BMI or obesity confers risk of MS exclusively through vitamin D deficiency, or whether other mechanisms related to obesity are involved.

Limitations of this study include the potential for inaccurate recall of body size, though we would expect this to bias results towards the null, as women that classify as overweight often underreport weight[52] and identify with a more favorable perception of body silhouette.[37] There is also the potential for selection bias, if controls participating in the study are healthier than nonparticipants with respect to body size/mass. Such differences could bias results away from the null. However, no association between BMI at time-of-interview or BMI in 30's and MS in females was detected. Additionally, 27% of the control population in our study classified as obese at time-of-interview, comparable to 26% of the California population using estimates of reported BMI in 2009.[53]

An additional limitation in our study was the small number of male MS cases and controls. Our recruitment pool is drawn from the eligible KPNC study participants meeting inclusion criteria and shows the female to male MS patient ratio in the overall KPNC membership is closer to 3:1; whereas, to date, our recruitment efforts indicate a 4:1 ratio. Females thus far have been more likely to participate, similar to what has been observed in other epidemiologic studies,[54] and extra efforts are currently being made to recruit more male participants. Stratified analyses of BMI and body size variables based on gender were performed in the current study, as well as separate power analyses by gender, which indicated that reduced power was available for males (Supplementary Data). We also studied white, non-Hispanic individuals, which may limit the external validity of our findings. Further, study of this relationship in African-Americans, as well as other populations, is needed.

While self-reported weight for a specific age has been shown to be valid,[25-28] we asked participants about the highest and lowest weight during a ten-year interval, which has not been specifically validated. However, it would appear that the same process that enables an individual to reliably recall their weight at a particular age would also enable an individual to review their weight at each age during a specified and limited interval, and be able to report fairly reliably their highest and lowest weights during that interval. Average BMI over the interval may be more accurate because it captures potential variability, rather than relying on recall of only one weight at a specific time. Weight is important to individuals in their 20's and the highest and lowest weights are likely to be remembered because of their salience in terms of body image. In the current study, the measure (average of highest and lowest BMI in the 20's) is significantly associated with MS, in line with what has been shown for recalled weight at age 18[14] and at age 20[15] in other studies. Many potential confounders have been assessed and controlled for, as described in the manuscript. We can think of no plausible unmeasured confounder that provides an alternative explanation for the association that we have found.

Finally, while a prospective cohort design can be used to fully establish temporality and minimize potential recall bias, a case-control design for a less prevalent disease such as MS, as described here, is essential to rigorously pursue a combined study of genetic and environmental risk factors with reasonable statistical power. Data were uniformly collected from all participants by trained interviewers or using standardized surveys. The observed association between BMI and MS in females within our study is very similar with respect to direction and magnitude to a previously published cohort study.[14] Similarity in results using both study designs underscores the utility of case-control studies to identify and model effects of multiple MS risk factors. Within our study specifically, we found nearly identical reporting of childhood and adolescent body size variables between cases with greater than 10 years of disease duration as compared to cases with less than 10 years of disease duration. We therefore do not expect results to vary between prevalent and incident MS cases.

In summary, the etiology of MS is very complex. The importance in MS susceptibility of both genetic and environmental factors, including obesity, is shown convincingly in our models. MS is a disease with high burden on society and quality of life, due primarily to disability.[2] Given the growing obesity epidemic in the U.S. and worldwide, these findings add to the increasing body of evidence for the involvement of obesity and related mechanisms in chronic diseases, including

MS. Our results demonstrate obesity as a potentially modifiable factor that could influence risk associated with developing MS in the population.

2.5 References

1. Sadovnick AD, Baird PA, Ward RH. Multiple sclerosis: updated risks for relatives. Am J Med Genet. 1988;29(3):533-541.
2. World Health Organization. Atlas: Multiple Sclerosis Resources in the World. Multiple Sclerosis International Federation 2008.
3. Confavreux C, Vukusic S, Moreau T, et al. Relapses and Progression of Disability in Multiple Sclerosis. N Eng J Med. 2000;343:1430-1438.
4. Cottrell DA, Kremenchutzky M, Rice GP, et al. The natural history of multiple sclerosis: a geographically based study. 5. The clinical features and natural history of primary progressive multiple sclerosis. Brain. 1999 Apr;122(4):625-39.
5. Sadovnick AD, Baird PA. The familial nature of multiple sclerosis: age-corrected empiric recurrence risks for children and siblings of patients. Neurology. 1988;38(6):990-991.
6. Robertson NP, Clayton D, Fraser M, et al. Clinical concordance in sibling pairs with multiple sclerosis. Neurology. 1996;47(2):347-352.
7. Favorova OO, Kulakova OG, Boiko AN. Multiple sclerosis as a polygenic disease: an update. Genetika. 2010 Mar;46(3):302-13.
8. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med. 2007 Aug 30;357(9):851-862.
9. Patsopoulos NA, Esposito F, Reischl J, et al. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Annals of neurology. 2011 Dec;70(6):897-912.
10. Sawcer S, Hellenthal G, Pirinen M, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011 Aug 11;476(7359):214-219.
11. Ascherio A, Munger K. Epidemiology of multiple sclerosis: from risk factors to prevention. Semin Neurol. 2008 Feb;28(1):17-28.
12. Ebers GC. Environmental factors and multiple sclerosis. Lancet Neurology. 2008 Mar;7(3):268-277.
13. Lauer K. Environmental risk factors in multiple sclerosis. Expert review of neurotherapeutics. 2010 Mar;10(3):421-440.
14. Munger KL, Chitnis T, Ascherio A. Body size and risk of MS in two cohorts of US women. Neurology. 2009 Nov 10;73(19):1543-1550.
15. Hedstrom 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. 2012 Sep;18(9):1334-1336.
16. Langer-Gould A, Brara SM, Beaber BE, et al. Childhood obesity and risk of pediatric multiple sclerosis and clinically isolated syndrome. Neurology. 2013 Feb 5;80(6):548-552.
17. Munger KL, Bentzen J, Lauresen B, et al. Childhood body mass index and multiple sclerosis risk: a long-term cohort study. Mult Scler. 2013 Apr 2 [Epub ahead of print].
18. Islam T, Gauderman WJ, Cozen W, et al. Childhood sun exposure influences risk of multiple sclerosis in monozygotic twins. Neurology. 2007 Jul 24;69(4):381-8.
19. Barcellos LF, Sawcer S, Ramsay PP, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Hum Mol Genet. 2006 Sep 15;15(18):2813-24.
20. Krieger N. Overcoming the absence of socioeconomic data in medical records: Validation and application of a census-based methodology. Am J Public Health. 1992;82:703-710.

21. McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001 Jul;50(1):121-7.
22. Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol. 2005 Dec;58(6):840-6.
23. Luzuriaga K, Sullivan JL. Infectious mononucleosis. N Engl J Med. 2010 May 27;362(21):1993-2000.
24. Handel AE, Williamson AJ, Disanto G, et al. An updated meta-analysis of risk of multiple sclerosis following infectious mononucleosis. PLoS One. 2010;5(9).
25. Stommel M, Schoenborn CA. Accuracy and usefulness of BMI measures based on self-reported weight and height: findings from the NHANES & NHIS 2001-2006. BMC Public Health 2009;9:421.
26. Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, Willet WC. Validity of self-reported waist and hip circumferences in men and women. Epidemiology 1990;1(6):466-473.
27. Park JY, Mitrou PN, Keogh RH, Luben RN, Wareham NJ, Khaw KT. Self-reported and measured anthropometric data and risk of colorectal cancer in the EPIC-Norfolk study. Int J Obes (Lond). 2012 Jan;36(1):107-118.
28. Troy LM, Hunter DJ, Manson JE, Colditz GA, Stampfer MJ, Willet WC. The validity of recalled weight among younger women. Int J Obes Relat Metab Disord 1995;19(8):570-572.
29. Stunkard AJ, Sorensen T, Schulsinger F. Use of the Danish Adoption Register for the study of obesity and thinness. Res Publ Assoc Res Nerv Ment Dis. 1983;60:115-120.
30. Barcellos LF, May SL, Ramsay PP, et al. High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. PLoS Genet. 2009 Oct;5(10):e1000696.
31. De Jager PL, Chibnik LB, Cui J, et al. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurology. 2009 Dec;8(12):1111-1119.
32. International Multiple Sclerosis Genetics Consortium (IMSGC). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013 Sept <Epub ahead of print>
33. Hedstrom AK, Sundqvist E, Baarnhielm M, et al. Smoking and two HLA genes interact to increase the risk for multiple sclerosis. Brain. 2011;134:653-664.
Ogden CL, Carroll MD, Kit BK, et al. Prevalence of obesity and trends in body mass index among US children and adolescents, 1999-2010. Jama. 2012 Feb 1;307(5):483-490.
34. Mokdad AH, Ford ES, Bowman BA, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. JAMA. 2003 Jan 1;289(1):76-79.
35. World Health Organization. Burden: mortality, morbidity and risk factors. Global status report on noncommunicable diseases 2010. 2011. Available online < http://www.who.int/nmh/publications/ncd_report_full_en.pdf>
36. Procaccini C, Carbone F, Galgani M, et al. Obesity and susceptibility to autoimmune diseases. Expert Rev Clin Immunol. 2011 May;7(3):287-294.
37. Tehard B, van Liere MJ, Com Nougue C, et al. Anthropometric measurements and body silhouette of women: validity and perception. J Am Diet Assoc. 2002 Dec;102(12):1779-1784.

38. Lim J, Iyer A, Liu L, et al. Diet-induced obesity, adipose inflammation, and metabolic dysfunction correlating with PAR2 expression are attenuated by PAR2 antagonism. FASEB. 2013;27:4757-4767.
39. Sbarbati A, Osculati F, Silvagni D, et al. Obesity and inflammation: evidence for an elementary lesion. Pediatrics. 2006 Jan;117(1):220-223.
40. Chu NF, Chang JB, Shieh SM. Plasma leptin, fatty acids, and tumor necrosis factor-receptor and insulin resistance in children. Obes Res. 2003 Apr;11(4):532-540.
41. Kanneganti TD, Dixit VD. Immunological complications of obesity. Nat Immunol. 2012 Aug;13(8):707-712.
42. Weinstock-Guttman B, Zivadinov R, Mahfooz N, et al. Serum lipid profiles are associated with disability and MRI outcomes in multiple sclerosis. J Inflamm. 2011; 8:127.
43. Ifergan I, Wosik K, Cayrol R, et al. Statins reduce human blood-brain barrier permeability and restrict leukocyte migration: Relevance to multiple sclerosis. Ann Neurol. 2006; 60:45-55.
44. Kau AL, Ahern PP, Griffin NW, et al. Human nutrition, the gut microbiome and the immune system. Nature. 2011 Jun 16;474(7351):327-336.
45. Lee YK, Menezes JS, Umesaki Y, et al. Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis. Proc Natl Acad Sci U S A. 2011 Mar 15;108 Suppl 1:4615-4622.
46. Tsai F, Coyle WJ. The Microbiome and Obesity: Is Obesity Linked to Our Gut Flora? Current Gastroenterology Reports. 2009;11:307-313.
47. Richardson AS, Dietz WH, Gordon-Larsen P. The association between childhood sexual and physical abuse with incident adult severe obesty across 13 years of the National Longitudinal Study of Adolescent Health. Pediatr Obes. 2013 Sep <Epub ahead of print>
48. Parikh SJ, Edelman M, Uwaifo GI, et al. The relationship between obesity and serum 1,25-dihydroxy vitamin D concentrations in healthy adults. J Clin Endocrinol Metab. 2004 Mar;89(3):1196-1199.
49. Smotkin-Tangorra M, Purushothaman R, Gupta A, et al. Prevalence of vitamin D insufficiency in obese children and adolescents. J Pediatr Endocrinol Metab. 2007 Jul;20(7):817-823.
50. Munger KL, Levin LI, Hollis BW, et al. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. JAMA. 2006 Dec 20;296(23):2832-2838.
51. Mowry EM, Waubant E, McCulloch CE, et al. Vitamin D status predicts new brain magnetic resonance imaging activity in multiple sclerosis. Ann Neurol. 2012 Aug;72(2):234-240.
52. Engstrom JL, Paterson SA, Doherty A, et al. Accuracy of self-reported height and weight in women: an integrative review of the literature. J Midwifery Womens Health. 2003 Sep-Oct;48(5):338-345.
53. Center for Disease Control. Prevalence and Trends Data: California – 2009. Behavioral Risk Factor Surveillance System. Office of Surveillance, Epidemiology, and Laboratory Services. <http://apps.nccd.cdc.gov/brfss/>
54. Galea S, Tracy M. Participation Rates in Epidemiologic Studies. Ann Epidemiol. 2007 Sept;17(9):643-653.

2.6 Tables and Figures

Table 1. Demographic and disease characteristics of KPNC MS cases and controls by gender[a]

Characteristic	Females			Males		
	MS Cases (N = 986)	Controls (N = 585)	P[b]	MS Cases (N = 249)	Controls (N = 112)	P[b]
Year of birth	1958 ± 8.88	1957 ± 8.24	0.02	1958 ± 9.03	1957 ± 8.50	0.13
Disease duration	12.07 ± 8.24	--	--	10.63 ± 8.31	--	--
Age at first symptom	31.18 ± 9.68	--	--	33.58 ± 9.09	--	--
Smoker			1.70×10^{-3}			0.35
Never	508 (51.63)	348 (59.79)		119 (47.79)	59 (53.15)	
Ever	476 (48.37)	234 (40.21)		130 (52.21)	52 (46.85)	
College graduate			1.44×10^{-4}			0.29
Yes	405 (41.08)	298 (50.94)		125 (50.20)	63 (56.25)	
No	581 (58.92)	287 (49.06)		124 (49.80)	49 (43.75)	
Infectious mononucleosis			2.80×10^{-8}			0.06
Yes	256 (26.31)	83 (14.35)		52 (21.05)	14 (12.61)	
No	717 (73.69)	498 (85.65)		195 (78.95)	97 (87.39)	
*HLA-DRB1**15:01 (N=1,708)			1.53×10^{-21}			6.50×10^{-4}
0	360 (41.33)	357 (67.61)		103 (49.05)	69 (69.70)	
1-2	511 (58.67)	171 (32.39)		107 (50.95)	30 (30.30)	
wGRS (N=1,666)	11.29 ± 0.70	11.11 ± 0.70	1.64×10^{-5}	11.33 ± 0.70	10.97 ± 0.70	3.82×10^{-5}
BMI at Time-of-Interview			0.46			0.05[c]
<18.5	26 (2.71)	9 (1.57)		1 (0.72)	0 (0.00)	
18.5-<21	121 (12.63)	67 (11.69)		5 (3.62)	0 (0.00)	
21-<23	159 (16.60)	95 (16.58)		19 (13.77)	5 (4.72)	
23-<25	137 (14.30)	100 (17.45)		27 (19.57)	21 (19.81)	
25-<27	135 (14.09)	73 (12.74)		26 (18.84)	30 (28.30)	
27-<30	115 (12.00)	76 (13.26)		30 (21.74)	24 (22.64)	
≥30	265 (27.66)	153 (26.70)		30 (21.74)	26 (24.53)	

[a] Table values are mean ± SD for continuous variables and n (column %) for categorical variables.
[b] *P*-value is for t-test (continuous variables) or χ^2 test (categorical variables) between MS Cases vs. Controls
[c] Fisher's exact test was utilized due to small cell counts

Table 2. Adjusted odds ratios (OR) demonstrating the association between childhood and adult body size and increased susceptibility to MS by gender

Characteristic	Females			Males		
	N (%MS)	Adjusted OR* (95% CI)	*P*-value	N (%MS)	Adjusted OR* (95% CI)	*P* -value
Body size at age 10						
Underweight	268 (60.07)	1.02 (0.77, 1.35)	0.90	73 (61.64)	0.72 (0.41, 1.26)	0.25
Just about right (Ref)	919 (59.85)	1.00	--	206 (68.93)	1.00	--
Little – Very overweight	339 (69.62)	1.50 (1.15, 1.97)	3.00×10^{-3}	77 (75.32)	1.30 (0.71, 2.38)	0.39
Body size at age 20						
Underweight	126 (49.21)	1.13 (0.77, 1.65)	0.54	36 (47.22)	0.57 (0.28, 1.17)	0.13
Just about right (Ref)	761 (44.94)	1.00	--	207 (59.90)	1.00	--
Little – Very overweight	229 (56.33)	1.60 (1.18, 2.16)	2.50×10^{-3}	30 (70.00)	1.50 (0.65, 3.46)	0.35
BMI in 20's						
<18.5	107 (47.66)	1.22 (0.79, 1.89)	0.36	9 (88.89)	3.79 (0.41, 34.64)	0.24
18.5-<21 (Ref)	371 (42.32)	1.00	--	32 (68.75)	1.00	--
21-<23	276 (49.64)	1.39 (1.02, 1.91)	0.04	68 (58.82)	0.67 (0.27, 1.63)	0.37
23-<25	150 (49.33)	1.38 (0.94, 2.03)	0.10	78 (52.56)	0.53 (0.22, 1.27)	0.15
25-<27	72 (55.56)	1.77 (1.06, 2.97)	0.03	40 (52.50)	0.51 (0.19, 1.37)	0.18
27-<30	52 (53.85)	1.63 (0.90, 2.94)	0.10	22 (63.64)	0.79 (0.25, 2.49)	0.68
≥30	53 (60.38)	2.15 (1.18, 3.92)	0.01	15 (66.67)	0.86 (0.23, 3.30)	0.83
		P-value trend	9.60×10^{-4}		*P-value trend*	0.33
BMI in 30's						
<18.5	25 (28.00)	1.22 (0.46, 3.25)	0.69	--	--	--
18.5-<21 (Ref)	168 (23.81)	1.00	--	13 (69.23)	1.00	--
21-<23	202 (24.26)	1.03 (0.63, 1.69)	0.91	24 (33.33)	0.25 (0.06, 1.10)	0.07
23-<25	110 (27.27)	1.21 (0.68, 2.15)	0.52	49 (38.78)	0.36 (0.09, 1.38)	0.14
25-<27	90 (18.89)	0.86 (0.44, 1.68)	0.66	35 (40.00)	0.37 (0.09, 1.47)	0.16
27-<30	62 (32.26)	1.87 (0.95, 3.69)	0.07	31 (29.03)	0.22 (0.05, 0.94)	0.04
≥30	72 (27.78)	1.69 (0.87, 3.30)	0.12	13 (23.08)	0.15 (0.03, 0.90)	0.04
		P-value trend	0.09		*P-value trend*	0.04

* Adjusted for year of birth, history of smoking, and college education (see Methods for details)

Table 3a. Multivariate models assessing the association between body size during childhood and increased susceptibility to MS in females[#]

	Conventional Covariates*		Conventional Covariates and Infectious Mononucleosis		Conventional Covariates, Infectious Mononucleosis, and Genotype	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Overall body size at age 10 (0=Just about right, 1= Little/Very overweight)	1.54 (1.15, 2.06)	3.50×10^{-3}	1.58 (1.18, 2.12)	2.10×10^{-3}	1.63 (1.21, 2.21)	1.50×10^{-3}
Year of Birth	1.01 (1.00, 1.02)	0.12	1.01 (1.00, 1.02)	0.14	1.01 (1.00, 1.03)	0.08
Smoker (0=no, 1=yes)	1.23 (0.98, 1.54)	0.08	1.23 (0.97, 1.54)	0.08	1.25 (1.03, 1.59)	0.06
College (0=yes, 1=no)	1.50 (1.20, 1.89)	4.4×10^{-4}	1.61 (1.27, 2.03)	6.10×10^{-5}	1.58 (1.24, 2.00)	2.10×10^{-4}
Infectious Mononucleosis (0=no, 1=yes)			2.26 (1.70, 3.00)	1.80×10^{-8}	2.17 (1.62, 2.91)	2.00×10^{-7}
*HLA-DRB1*15:01* positive (0=no, 1=yes)					2.90 (2.29, 3.68)	1.70×10^{-18}
wGRS					1.41 (1.19, 1.68)	7.20×10^{-5}

Analyses restricted to individuals with complete genetic information and age of onset > 15 years (N=1327)

* Conventional covariates include year of birth, history of smoking, and college education.

Table 3b. Multivariate models assessing the association between body size in 20's and increased susceptibility to MS in females[#]

	Conventional Covariates*		Conventional Covariates and Infectious Mononucleosis		Conventional Covariates, Infectious Mononucleosis, and Genotype	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Overall body size at age 20 (0=Just about right, 1= Little/Very overweight)	1.62 (1.17, 2.24)	3.80×10^{-3}	1.62 (1.17, 2.25)	4.00×10^{-3}	1.70 (1.21, 2.39)	2.50×10^{-3}
Year of Birth	0.97 (0.96, 0.99)	2.90×10^{-3}	0.97 (0.96, 0.99)	3.30×10^{-3}	0.98 (0.96, 1.00)	0.02
Smoker (0=no, 1=yes)	1.11 (0.86, 1.43)	0.44	1.11 (0.86, 1.44)	0.44	1.21 (0.88, 1.51)	0.30
College (0=yes, 1=no)	1.50 (1.16, 1.94)	2.10×10^{-3}	1.59 (1.22, 2.07)	5.20×10^{-4}	1.53 (1.15, 1.98)	3.20×10^{-3}
Infectious Mononucleosis (0=no, 1=yes)			2.01 (1.46, 2.76)	1.80×10^{-5}	2.34 (1.45, 2.79)	3.00×10^{-5}
*HLA-DRB1*15:01* positive (0=no, 1=yes)					2.83 (2.17, 3.72)	3.40×10^{-14}
wGRS					1.43 (1.18, 1.73)	2.30×10^{-4}

Analyses restricted to individuals with complete genetic information and age of onset $\geq$ 30 years (N=981)

* Conventional covariates include year of birth, history of smoking, and college education.

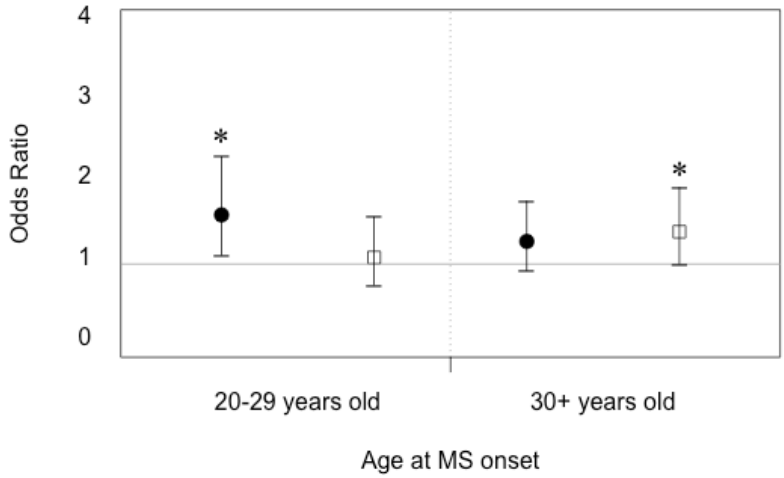
Table 3c. Multivariate models assessing the association between mean BMI in 20's and increased susceptibility to MS in females[#]

	Conventional Covariates*		Conventional Covariates and Infectious Mononucleosis		Conventional Covariates, Infectious Mononucleosis, and Genotype	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Mean BMI in 20's (0= 18.5-<21 kg/m^2 1= ≥ 30 kg/m^2)	2.51 (1.29, 4.87)	6.50×10^{-3}	2.67 (1.36, 5.22)	4.20×10^{-3}	2.98 (1.49, 5.94)	2.00×10^{-3}
Year of Birth	0.97 (0.95, 0.99)	7.20×10^{-4}	0.97 (0.95, 0.99)	6.50×10^{-4}	0.97 (0.96, 0.99)	5.20×10^{-3}
Smoker (0=no, 1=yes)	1.14 (0.88, 1.49)	0.32	1.15 (0.88, 1.50)	0.29	1.20 (0.91, 1.58)	0.20
College (0=yes, 1=no)	1.46 (1.12, 1.91)	4.90×10^{-3}	1.54 (1.18, 2.02)	1.60×10^{-3}	1.45 (1.09, 1.92)	9.60×10^{-3}
Infectious Mononucleosis (0=no, 1=yes)			2.08 (1.50, 2.88)	1.10×10^{-5}	2.12 (1.51, 2.96)	1.30×10^{-5}
*HLA-DRB1*15:01* positive (0=no, 1=yes)					2.82 (2.14, 3.71)	2.00×10^{-13}
wGRS					1.48 (1.22, 1.80)	8.30×10^{-5}

Analyses restricted to individuals with complete genetic information and age of onset ≥ 30 years (N=952)

* Conventional covariates include year of birth, history of smoking, and college education.

Figure 1.



Multivariate adjusted odds ratios and 95% CIs for body size/mass and MS susceptibility in females at different time periods. Stratified odds ratios are presented for MS cases with age of onset between 20-29 years of age, and those with age at onset at 30 years old or later. Black circles reflect odds of MS amongst females reporting being a little/very overweight at age 10. White squares demonstrate odds of MS amongst females reporting being a little/very overweight at age 20. Asterisk indicates *P*-value <0.05.

Supplementary information:

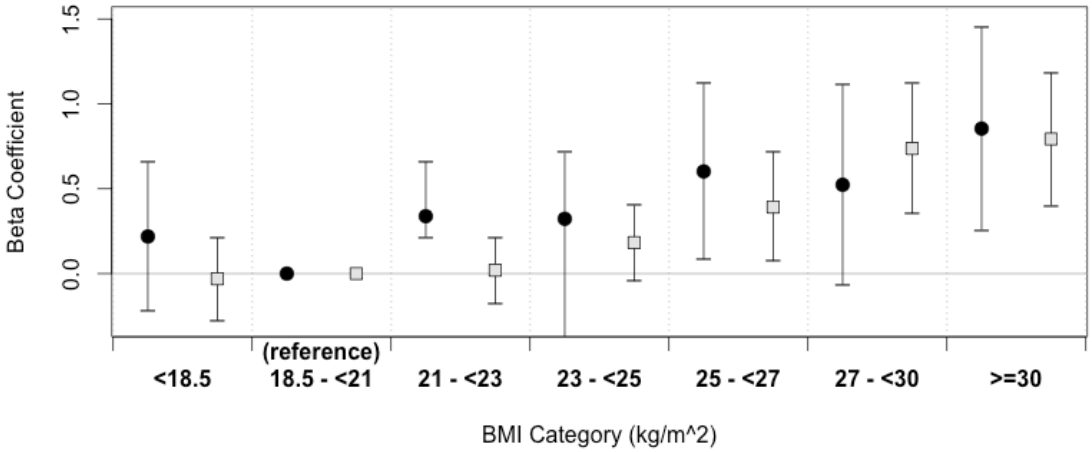
Analyses were conducted to determine available power to detect marginal associations between BMI and MS. ORs ranging from 0.1 to 3.0 were examined, assuming a two-sided type-1 error of 5% (α=0.05). Results indicated that our study in females had sufficient power (>75%) to detect OR <0.6 and >1.5 for exposures with a prevalence ≥13.0%, which reflects the prevalence of overweight/obese female controls from KPNC. In males, sufficient power was present to detect OR <0.5 and >1.9 for exposures with a prevalence ≥ 29.0%, which reflects the prevalence of overweight/obese male controls from KPNC.

Supplementary Table 1. Unadjusted and adjusted odds ratios demonstrating the association between childhood and adult body size as reported by silhouette identification and increased susceptibility to MS in KPNC females

Characteristic	N (%MS)	Unadjusted OR (95% CI)	*P*-value	Adjusted OR* (95% CI)	*P*-value
Body Size at Baseline					
1	14 (92.9)	6.93 (0.88, 54.38)	0.06	7.56 (0.96, 59.5)	0.06
2	59 (69.5)	1.21 (0.64, 2.31)	0.55	1.13 (0.59, 2.15)	0.72
3	161 (65.2)	1.00	--	1.00	--
4	271 (63.1)	0.91 (0.61, 1.37)	0.66	0.87 (0.57, 1.31)	0.49
5	325 (58.5)	0.75 (0.51, 1.11)	0.15	0.73 (0.49, 1.08)	0.12
≥6	322 (67.1)	1.09 (0.73, 1.62)	0.68	1.05 (0.70, 1.57)	0.82
		P-value trend	0.57	*P-value trend*	0.45
Body Size at 10					
1	255 (63.1)	1.20 (0.82, 1.75)	0.36	1.18 (0.80, 1.74)	0.40
2	294 (59.2)	1.01 (0.70, 1.46)	0.95	1.02 (0.70, 1.47)	0.94
3	197 (58.9)	1.00	--	1.00	--
4	176 (65.3)	1.32 (0.86, 2.00)	0.20	1.32 (0.86, 2.01)	0.20
5	153 (69.9)	1.62 (1.04, 2.54)	0.03	1.59 (1.01, 2.49)	0.04
≥6	54 (74.1)	2.00 (1.02, 3.90)	0.04	1.91 (0.97, 3.76)	0.06
		P-value trend	0.01	*P-value trend*	0.02
Body Size at 20					
1	37 (54.0)	1.57 (0.79, 3.12)	0.20	1.46 (0.73, 2.92)	0.28
2	174 (47.1)	1.23 (0.83, 1.70)	0.30	1.17 (0.79, 1.72)	0.43
3	240 (42.9)	1.00	--	1.00	--
4	190 (46.3)	1.17 (0.80, 1.70)	0.42	1.21 (0.82, 1.77)	0.33
5	99 (53.5)	1.54 (0.97, 2.46)	0.07	1.58 (0.99, 2.52)	0.06
≥6	46 (47.8)	1.27 (0.68, 2.37)	0.45	1.28 (0.68, 2.40)	0.44
		P-value trend	0.14	*P-value trend*	0.10
Body Size at 30					
1	4 (25.0)	1.59 (0.26, 9.82)	0.62	1.57 (0.25, 9.95)	0.63
2	65 (20.0)	0.73 (0.38, 1.41)	0.35	0.72 (0.37, 1.42)	0.35
3	154 (27.3)	1.00	--	1.00	--
4	163 (19.6)	0.71 (0.43, 1.16)	0.17	0.76 (0.46, 1.27)	0.30
5	91 (19.8)	0.72 (0.40, 1.29)	0.27	0.81 (0.44, 1.49)	0.50
≥6	60 (25.0)	1.01 (0.53, 1.90)	0.98	1.52 (0.77, 2.99)	0.23
		P-value trend	0.40	*P-value trend*	0.93

* Adjusted for age, history of smoking, and college education (see Methods for details)

Supplementary Figure 1.



Multivariate adjusted beta coefficients for BMI category and MS susceptibility in females. Black circles represent values in the KPNC dataset for BMI category in 20's; grey squares represent values in the EIMS dataset for BMI category at age 20.

Chapter 3

Genetic variants associated with body mass index demonstrate a causal effect on multiple sclerosis susceptibility

3.1 Background

MS is a severe and complex disease of the central nervous system resulting in significant disability and decreased quality of life. The disease is characterized as an immune-mediated, demyelinating disorder with widespread axonal degeneration occurring throughout the disease process.[1] Strong evidence supports the contribution of both genetic and environmental factors to MS disease susceptibility.[2] Substantial progress has been made towards the identification of several MS genetic risk factors including the *HLA-DRB1*15:01* allele within the major histocompatibility complex (MHC),[3, 4] and 110 variants outside of the MHC in individuals of European ancestry[4]. Environmental risk factors associated with MS susceptibility include exposure to tobacco smoke, Epstein-Barr Virus (EBV) infection, and low levels of vitamin D.[5]

Recently, obesity has also emerged as a significant risk factor for MS. Association between MS and body mass index (BMI) at age 18,[6] age 20,[7] and during one's twenties[8] has been observed, where individuals having a BMI $\geq$ 30 kg/m^2 demonstrated greater than a twofold increased risk of MS compared to those at normal weight (18.5-<21 kg/m^2). The relationship was also confirmed using retrospective assessments of body size, with the strongest association at age 25.[9] Additionally, childhood obesity and risk of both pediatric[10] and later onset[11] MS was reported. Importantly, findings in one study remained significant after controlling for established genetic and environmental risk factors for the first time.[8] Research has also indicated that HLA genes interact with BMI during adolescence to increase the risk of MS.[12] Although studies have hypothesized various mechanisms mediate the association between obesity and MS, a causal relationship remains to be confirmed. Observational findings are not able to exclude the possibility that weight change preceding diagnosis may be associated with a prodromal period of disease; that is, a decrease in physical activity due to MS symptoms before a diagnosis is made may lead to increased obesity, and thus bias findings.

The genetic basis of obesity has been demonstrated by studies revealing multiple variants throughout the genome associated with obesity-related traits. A recent genome-wide association study (GWAS) identified 97 significant loci associated with BMI.[13] To date, no study has examined the relationship between genes associated with BMI and MS. Using genetic predictors of an exposure of interest as independent variables in an observational study corresponds with a "Mendelian Randomization" study design that avoids bias from reverse causation and many potential confounders, therefore strengthening causal inferences with complex and potentially confounded exposures such as BMI. Additionally, understanding the downstream effects of genes associated with BMI has the potential to aid in the understanding of MS disease pathogenesis.

We applied instrumental variable analysis in a Mendelian Randomization (MR) framework to estimate the causal relationship between BMI and MS susceptibility using a BMI genetic risk

score (GRS) compromising 97 variants[13] in two populations. We further examined whether any variants appeared to directly influence MS susceptibility via mechanisms unrelated to BMI, and if evidence of protein-protein interaction networks of MS and BMI variants could be identified using bioinformatics analyses.

3.2 Methods

KPNC Participants

MS cases and controls were recruited from members of Kaiser Permanente Medical Care Plan, Northern California Region (KPNC). KPNC is an integrated health services delivery system with a membership of 3.2 million that comprises about 25-30% of the population of a 22 county service area and is the largest healthcare provider in northern California. Membership is objectively representative of the general population; however, persons in impoverished neighborhoods are underrepresented.[14] The KPNC membership is stable with 64% of all members, and over 72% of those aged 40 or more years, maintaining membership for five years or more. Individuals with chronic conditions such as MS have historically been more likely to remain members.

This study was restricted to self-identified white (non-Hispanic) race/ethnicity, the population with the highest prevalence of MS. Eligible KPNC cases were defined as: individuals with a diagnosis of MS by a neurologist (ICD9 code 340.xx; 94.7% of cases had at least two MS diagnoses by a neurologist), age of 18 through 69 years, and membership in KPNC at initial contact. The treating neurologist was contacted for approval to contact each case as a potential MS study participant. A total of 3,293 potential MS cases were reviewed by KPNC neurologists, who approved contact with 2,823 (86%) at the time of the data freeze (August 2014). Diagnoses were validated utilizing electronic health record (EHR) review and published diagnostic criteria.[15, 16]

Controls were KPNC current members without a diagnosis of MS or related condition (optic neuritis, transverse myelitis, or demyelination disease; ICD9 codes: 340, 341.0, 341.1, 341.2, 341.20, 341.21, 341.22, 341.8, 341.9, 377.3, 377.30, 377.39, and 328.82) confirmed through electronic records, and white (non-Hispanic) race/ethnicity. Controls were matched to cases on age, gender, and zipcode. Potential study participants were contacted by mail with a follow-up phone call to explain the study and procedures. The participation rate was approximately 80% for cases and 66% for controls. Genetic data were available for approximately 80% of study participants.

Additional controls were individuals in the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort who participated in the KPNC Research Program on Genes, Environment, and Health (RPGEH). The RPGEH was established to research genetic and environmental influences on health and disease and is described in more detail elsewhere (dbGaP phs000674.v2.p2). GERA cohort members completed a broad written consent and provided a saliva sample for DNA extraction.[17] A total of 110,266 participant samples were successfully genotyped; approximately 77% returned completed new consent forms for placement in dbGaP (NIH), resulting in a final sample size of 78,486 participants. From these participants, we selected a subset of 12,605 self-reported non-Hispanic white individuals without evidence of MS

in EHR matched to MS cases on gender and age (+/- 2 years) at a 10:1 ratio. Study protocols were approved by the Institutional Review Boards (IRB) of KPNC and the University of California, Berkeley.

Swedish participants

Data were collected from two population based case-control studies on incident (EIMS study) and prevalent (GEMS study) MS patients. The EIMS study inclusion criteria were: age 16-70 years, diagnosed MS according to the to the McDonald criteria[15,16] within two years, and ability to understand the Swedish language. GEMS study participants were identified from the Swedish National MS registry, fulfilled the McDonald criteria,[15,16] and were recruited during 2009-2011. For both studies, controls were randomly chosen from the population register and matched to MS patients by sex, age at inclusion in the study, and region of residence. Two controls were matched to each case in the EIMS study and one control per case in the GEMS study. All participants in the EIMS study were distinct from those in the GEMS study. Ethical approval for both studies was obtained from the Regional Ethical Review Board in Stockholm at Karolinska Institutet and participants provided written informed consent. Details of the study design have been described elsewhere.[7,12] The participation rate in the EIMS study was 92% for cases and 67% for controls, and in the GEMS study 82% for cases and 66% for controls. Genotyping data were available for 75% of EIMS and 91% of GEMS participants.

KPNC Exposure Assessment

KPNC study participants completed a computer-assisted telephone interview (CATI) administered by trained staff interviewers and comprised of questions related to various events and exposures as described elsewhere.[8] GERA controls completed a survey consisting of questions related to health behaviors, sociodemographic information, and diagnoses (dbGaP phs00674.v1.p1).

KPNC study participants reported their highest and lowest (non-pregnancy) weight during their 20's. Mean weight of each KPNC participant during their 20's was calculated by averaging the highest and lowest weight reported. GERA controls reported their weight at age 18. Each participant's BMI was calculated by dividing weight in kilograms (or mean weight for KPNC) by height reported at time-of-interview in meters squared.

Exposure assessment in the Swedish studies was done through an extensive questionnaire that participants completed at home. Incomplete questionnaires were completed by mail or telephone. The questionnaire covered demographic and environmental/lifestyle factors, including current height and weight at age 20. The details have been described elsewhere.[7,12]

Each participant's BMI was calculated by dividing weight in kilograms (or mean weight for KPNC) by height reported at time-of-interview in meters squared. KPNC and Swedish participants provided blood or saliva samples for genotyping. Details on platforms and quality control can be found in Supplementary Information. Analyses were carried out with genotypes or imputed genotype dosages.

Statistical Analyses

A weighted genetic risk score (wGRS) of MS risk variants was calculated for each individual that weights risk alleles by the logarithm of the odds ratio for each of the 110 non-HLA MS susceptibility loci identified through the most recent MS GWAS.[4] The wGRS was calculated by multiplying the number of risk alleles for each locus by the weight for that variant, and then taking the sum across the 110 loci. One SNP was missing for KPNC (rs201202118), and two SNPs were missing for the Swedish studies (rs2028597, rs6874308).

The BMI GRS was calculated by multiplying the number of risk alleles for each BMI-related locus by the weight (defined as the beta coefficient from the BMI GWAS[13]) for that variant and then taking the sum across the 97 loci. One SNP was missing for the Swedish dataset (rs2245368). To correct for the fact that the Locke et al GWAS was based on a residualized BMI transformed into standard deviation units and express this weight in terms of BMI units, we multiplied by a constant of 4.95. This value was estimated using the regression amongst European-descent individuals in the Health and Retirement Study as reported in the GWAS.[13] With this transformation, each unit change in the weighted BMI GRS corresponds with an anticipated one unit change in BMI, allowing for direct interpretation of the effect estimates as the estimated effect of a unit increase in the genetically predicted BMI on the odds of MS (BMI GRS range = 8.64 - 14.52 in KPNC, and 8.41 – 14.33 in EIMS/GEMS). It should be noted that none of the direct effect variants or genes, nor any of the variants or genes from the BMI GRS, overlap with the 110 non-HLA MS risk loci.

After quality control and removal of population outliers, a total of 2,163 individuals from KPNC (1,104 cases, 804 controls) with genetic data were available as well as an additional 9,732 controls from the GERA study for a total of 1,104 cases and 10,536 controls. Data for 6,335 cases and 5,762 controls was available from the Swedish studies (EIMS and GEMS).

Demographic differences between cases and controls were compared using X^2 tests and independent sample t-tests where appropriate. Linear regression was used to demonstrate the association of each GRS with BMI during young adulthood (in one's 20's for KPNC, at age 18 for GERA, and at age 20 for Sweden), and to test the assumption that each GRS is not associated with confounding factors. MR analysis, in this case a separate-sample instrumental variable analysis using weights[13] as described above, was performed by regressing MS cases status on the BMI GRS. All analyses were controlled for year of birth, ever smoking, college education, *HLA-DRB1*15:01*, wGRS of non-HLA MS risk variants, and ancestry as derived from MDS components. Swedish analyses were additionally controlled for region of residency and study type (EIMS vs. GEMS). Meta-analysis assuming random effects was performed.

In addition, we evaluated evidence that any of the BMI variants had a direct effect on MS susceptibility, implying a violation of the MR assumption that there is no direct effect of the instrument on the outcome.[18] Direct effects were analyzed using regression-based mediation analysis proposed by Valeri and VanderWeele (2013) to estimate the controlled direct effect (CDE) for changes in exposure level.[18] Analyses examined 97 BMI variants, measuring the effect of having no increasing BMI risk alleles ($a = 0$) versus having two risk alleles ($a = 1$) at each locus on MS status (case/control). The mediator was specified as BMI during young adulthood and set to 22.0 kg/m^2, or the mean of the "normal" healthy BMI range as defined by the World Health Organization for each of the 97 BMI variants. Models were adjusted for sex,

year of birth, ancestry, smoking, wGRS of non-HLA MS risk variants and number of *DRB1*15:01* alleles. The Swedish data were additionally controlled for region of residency and study type (EIMS vs. GEMS). After bootstrapping analyses with 100 replications, we used estimates from both datasets (KPNC and Sweden) to conduct a random-effects meta-analysis.

The Disease Association Protein-Protein Link Evaluator (DAPPLE, v2.0, Broad Institute) was utilized for bioinformatics analysis to examine whether protein-protein interactions exist between established genome-wide significant gene regions associated with both BMI (n=97) and MS disease susceptibility (n=110). Unweighted gene sub-scores based on resulting protein networks were derived by summing the number of risk alleles across each representative gene within each network. Logistic regression was used to test for association between each network sub-score and MS susceptibility, controlling for smoking, education, year of birth, genetic ancestry, and *HLA-DRB1*15:01*. Analyses using the Swedish dataset additionally controlled for study type. The following specifications were made in DAPPLE: genome assembly (Release 23, Hg 19, HapMap); number of permutations: 1000; common interactor binding degree: 2; and gene regulatory region: 50kb up/downstream. We selected to use the nearest gene, and did not specify any genes. The genes to which each variant was mapped by DAPPLE are listed in Supplementary Table I.

Odds ratios with 95% CIs were estimated. Analyses were conducted in PLINK, STATA, and R. This study was focused on a single hypothesis established a priori: BMI is causally associated with MS as represented by a BMI GRS. Therefore, we report 95% confidence intervals and use an α=0.05 threshold for statistical significance.

3.3 Results

There were significant differences between cases and controls with respect to smoking, college graduation, *HLA-DRB1*15:01* status, and wGRS in KPNC and Swedish studies (Table 1). BMI during young adulthood was significantly higher for cases than controls, as was the BMI GRS in both datasets. The BMI GRS was associated with college education and wGRS in KPNC, and smoking in both populations. The association between the BMI GRS and BMI in young adulthood was more pronounced in males compared to females (Table 2). For every unit increase in the GRS, BMI in young adulthood increased on average 0.64 units (kg/m^2) in KPNC and 0.54 units (kg/m^2) in the Swedish dataset.

The BMI GRS significantly predicted odds of MS after controlling for sex, year of birth, ancestry, smoking, wGRS and number of *HLA-DRB1*15:01* alleles in KPNC (OR = 1.13, 95% CI 1.04, 1.22; Table 3, Figure 1). Similar results were found for Sweden after controlling for the same covariates as well as study type (OR = 1.09, 95% CI 1.03, 1.15), as well as the meta-analysis of both study findings (OR=1.10, 95% CI 1.05, 1.15). No evidence of heterogeneity between populations was observed (I^2=0.0%, heterogeneity X^2 *P*=0.47).

Results stratified by sex showed a significant effect in females (OR = 1.14, 95% CI 1.04, 1.25), but not males (OR=1.11, 95% CI 0.92, 1.33) in KPNC, as well as Sweden (OR_{female}= 1.09, 95% CI 1.03, 1.16 and OR_{male} =1.08, 95% CI 0.97, 1.20); however, the wide CIs suggest this is consistent with chance given the smaller sample size of male participants.

The meta-analysis results of the direct effect estimates demonstrated that five variants associated with increased BMI in the literature exhibited a significant controlled direct effect on MS susceptibility after adjusting for covariates (Table 4). Four were positively associated with MS, and one was inversely associated with MS.

Given the potential violation of assumptions, i.e. that there is no direct effect of the instrument on the outcome, the MR analysis was re-estimated using a BMI GRS excluding the five variants for which we found evidence of a possible direct effect on MS. The findings demonstrated a consistent significant association between the revised 92 variant BMI GRS and MS after controlling for covariates in both KPNC and Sweden (OR = 1.13, 95% CI 1.04, 1.23 and OR = 1.09, 95% CI 1.03, 1.15, respectively).

We also adjusted the MR analysis of 92 variants for BMI during young adulthood to examine evidence that the GRS had effects on MS not mediated by BMI. Adjustment attenuated our findings and reduced significance of the association in the KPNC (OR = 1.05, 95% CI 0.97, 1.15) and Swedish dataset (OR = 1.06, 95% CI 1.00, 1.12).

Validating additional assumptions
In order to test the MR model assumptions for the BMI GRS, we conducted overidentification tests to evaluate the null hypothesis that effect estimates from multiple IVs are identical (21). The 92 variants without a direct effect on the outcome were randomly split into five separate instruments. Estimates suggested the same direction of causal effect, with ORs ranging from 1.04 to 1.22 in KPNC, and 1.01 to 1.25 in Sweden.

Network Analysis
Results revealed significant evidence for direct and indirect protein-protein interactions; of 99 total direct interactions observed, 22 were between obesity and MS genes (Supplementary Table II). The interactions clustered into 13 networks (Figure 2). Five of were significantly associated with MS susceptibility in both the KPNC and Swedish dataset after controlling for covariates (P<0.05) (Table IV), and three of these included both obesity and MS genes.

3.4 Discussion

This study was the first to examine the relationship between BMI and MS using Mendelian randomization. We also present novel results suggesting that five variants previously established to predict BMI may have direct effects on MS susceptibility. Our results suggest a causal association between higher BMI and MS susceptibility. Sub-scores based on protein-protein interaction pathways between BMI and MS risk variants also demonstrated significant associations related to MS susceptibility, suggesting that specific networks may contribute to disease onset.

The most widely studied obesity-related genetic variant is the fat mass and obesity associated gene (*FTO*).[20,21] Variants in *FTO* have been found to significantly increase the risk of various cancers, such as breast, prostate, and endometrial cancer,[22-24] as well as Alzheimer's disease, dementia,[25] reduced brain volume in healthy elderly individuals,[26] and cognitive decline in

healthy adults.[27] Although one recent study found that the *FTO* risk allele was associated with significantly increased homocysteine levels in MS cases compared to controls,[28] no study has previously examined the relationship between BMI genes and MS susceptibility. Our study found that *FTO* alone slightly increased the risk of MS, and that there was some evidence of a direct effect of *FTO* on MS susceptibility.

Since the discovery of *FTO*, additional genes have been cited in the literature as being associated with obesity-related traits. The recent study by Locke et al. (2015) reported 97 variants in the largest GWAS meta-analysis of BMI to date, of which 56 were novel loci. Previous studies using a subset of these variants improved prediction of BMI and obesity beyond demographic, geographic, and socioeconomic status information.[29] This subset of variants has also demonstrated a significant influence on BMI during childhood, adolescence, and adulthood in a longitudinal cohort.[30] Thus, the variants seem to confer a life course of obesity risk rather than at one time-point. Our study showed that a score including the cumulative effects of all recent GWAS variants was significantly associated with MS susceptibility. Effect estimates using the additional loci were statistically consistent with effect estimates based on *FTO* only, as would be expected if all these variants influenced MS via a common pathway, e.g., BMI. This perspective is also supported by our over-identification tests, which found no statistically significant difference in effects across five arbitrary groupings. We also found very similar effect estimates when we used all 97 variants or excluded five with some evidence of a direct effect. These three lines of evidence suggest that although some of the genetic variants may have small direct effects on MS, there is likely to be a common pathway to MS mediated by BMI.

There are several hypotheses linking obesity and autoimmune diseases, including MS. T-helper 17 cells, which secrete IL-17, have recently been implicated in the pathogenesis of autoimmune disease, and obesity may predispose induction of TH17 cells via an IL-6 dependent process leading to exacerbation of inflammatory diseases such as MS.[31] The intestinal immune response has also been hypothesized to explain the association between obesity and MS, as Th17/Treg imbalance may lead to alteration of intestinal microbiome.[32,33] It has also been shown that vitamin D deficiency is prevalent amongst obese individuals.[34] Given that vitamin D regulates immune response and has been shown to increase Treg cells and inhibit Th1 and Th17 differentiation,[35] this may have implications on MS susceptibility. Additionally, white adipose tissue has been identified as an essential endocrine organ that secretes adipokines (e.g. IL-6, TNF-alpha, leptin, adiponectin), which are involved in immune and inflammatory processes and contribute to the low-grade inflammatory state present in obese individuals.[36] Certain adiposity genes, some of which overlap with genes in the BMI GRS used in this study, have been shown to be associated with age-of-menarche,[37] and adipokines such as leptin are up-regulated by ovarian sex steroids.[38] In fact, higher levels of leptin have been found in females compared to males,[39] and leptin-deficient mice have been shown to be resistant to experimental autoimmune encephalomyelitis (EAE), an animal model of MS.[39] Lastly, a recent study demonstrated genetic evidence for overlap between motor deficits, obesity and neurological disorders.[40] More research needed to identify the biological mechanisms mediating the association between obesity in young adulthood and risk of MS, and examine if and how specific BMI genes may relate to leptin and whether that may explain the predominance of MS in females.

A major strength of this study is the large sample size and ability to have power to demonstrate

an association using a relatively weak instrument (BMI GRS R^2 = 1.5% [Sweden] - 2.3% [KPNC]), which is in accordance with the GWAS identifying the 97 variants (R^2 = 2.7%).[13] Additionally, we were able to replicate our findings in a second dataset. We conducted rigorous quality control to account for population stratification, and we able to model both genetic and environmental risk factors associated with disease susceptibility. Further, strengths include the fact that cases and controls were sampled from the same source population (i.e. Kaiser Permanente, Northern California patient population), and we used the most recent data to build a BMI GRS of 97 variants. Lastly, we examined potential biological pathways underlying the relationship between BMI and MS by utilizing a bioinformatics database. Understanding these mechanisms could contribute to our understanding of obesity and MS, as well as other autoimmune and neurological diseases.

As with any efforts to make a causal inference from observational data, MR analysis involves many assumptions, we have addressed them as best as possible. The GRS is a valid instrumental variable for the effect of BMI on MS if: a) it is associated with BMI; b) it is independent of measured or unmeasured confounders; and c) it can only influence that outcome via the causal effect of the exposure.[41] We were able to meet most model assumptions by utilizing a BMI GRS established to be associated with BMI in an independent population through a large GWAS, testing whether the BMI GRS is independent of measured confounders, and conducting a direct effect analysis to ensure that the BMI GRS did not contain any variants independently associated with MS. We additionally validated our BMI GRS by conducting overidentification tests and adjusting our model for self-reported BMI. However, associations of genetic variants with unmeasured or unknown confounders cannot be ruled out, leaving one assumption not fully testable.[41]

Our study included non-Hispanic whites, which limits the generalizability of our findings. Additional limitations of our study include a relatively small male sample size, assumption of linearity, and possible pleiotropic effects of BMI GRS on MS (i.e. genes may influence phenotypes other than BMI that are associated with an increased risk of MS). Further, the life course specificity of the BMI GRS remains to be fully understood. Limitations of direct effects analysis include reliance on self-reported weight/height to calculate BMI, which may bias the mediation results. We also used two slightly different assessments to calculate self-reported BMI during young adulthood for the KPNC (weight in one's 20's) and GERA (weight at age 18) cohorts; however, repeated sampling procedures in the GERA dataset indicated no significant deviations in BMI between the two control samples. Additional studies should aim to replicate findings, specifically in populations of other race/ethnicities, and better examine how specific BMI-related variants may influence MS susceptibility and severity.

In conclusion, we found BMI to be associated with MS susceptibility by utilizing a BMI GRS of 97 variants. While BMI-related variants demonstrate indirect effects on MS susceptibility through their association with increased BMI, we also found that certain variants may directly influence MS susceptibility via independent mechanisms. Our results are consistent with previous studies demonstrating an association between obesity and MS, and confirm a causal association. These findings aid in the understanding of the complex relationship between genetics, BMI, and the MS disease process.

3.5 References

1. Su KG, Banker G, Bourdette D, et al. Axonal degeneration in multiple sclerosis: the mitochondrial hypothesis. Curr Neurol Neurosci Rep. 2009 Sep;9(5):411-7.
2. Favorova OO, Kulakova OG, Boiko AN. [Multiple sclerosis as a polygenic disease: an update]. Genetika. 2010 Mar;46(3):302-13.
3. Barcellos LF, Sawcer S, Ramsay PP, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Human molecular genetics. 2006 Sep 15;15(18):2813-24.
4. Beecham AH, Patsopoulos NA, Xifara DK, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013 Nov;45(11):1353-60.
5. Ascherio A, Munger K. Epidemiology of multiple sclerosis: from risk factors to prevention. Seminars in neurology. 2008 Feb;28(1):17-28.
6. Munger KL, Chitnis T, Ascherio A. Body size and risk of MS in two cohorts of US women. Neurology. 2009 Nov 10;73(19):1543-50.
7. Hedstrom 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. 2012 Sep;18(9):1334-6.
8. Gianfrancesco MA, Acuna B, Shen L, et al. Obesity during childhood and adolescence increases susceptibility to multiple sclerosis after accounting for established genetic and environmental risk factors. Obes Res Clin Pract. 2014 Sep-Oct;8(5):e435-47.
9. Wesnes K, Riise T, Casetta I, et al. Body size and the risk of multiple sclerosis in Norway and Italy: The EnvIMS study. Mult Scler. 2014 Sep 2.
10. Langer-Gould A, Brara SM, Beaber BE, et al. Childhood obesity and risk of pediatric multiple sclerosis and clinically isolated syndrome. Neurology. 2013 Feb 5;80(6):548-52.
11. Munger KL, Bentzen J, Laursen B, et al. Childhood body mass index and multiple sclerosis risk: a long-term cohort study. Mult Scler. 2013 Sep;19(10):1323-9.
12. Hedstrom AK, Lima Bomfim I, Barcellos L, et al. Interaction between adolescent obesity and HLA risk genes in the etiology of multiple sclerosis. Neurology. 2014 Mar 11;82(10):865-72.
13. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015 Feb 12;518(7538):197-206.
14. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. Am J Public Health. 1992 May;82(5):703-10.
15. McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Annals of neurology. 2001 Jul;50(1):121-7.
16. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of neurology. 2011 Feb;69(2):292-302.
17. Kvale MN, Hesselson S, Hoffmann TJ, et al. Genotyping Informatics and Quality Control for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. Genetics. 2015 Jun 19.
18. Glymour MM, Tchetgen Tchetgen EJ, Robins JM. Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. Am J Epidemiol. 2012 Feb 15;175(4):332-9.

19. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods. 2013 Jun;18(2):137-50.
20. Scuteri A, Sanna S, Chen WM, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet. 2007 Jul;3(7):e115.
21. Loos RJ, Bouchard C. FTO: the first gene contributing to common forms of human obesity. Obes Rev. 2008 May;9(3):246-50.
22. Kaklamani V, Yi N, Sadim M, et al. The role of the fat mass and obesity associated gene (FTO) in breast cancer risk. BMC Med Genet. 2011;12:52.
23. Lewis SJ, Murad A, Chen L, et al. Associations between an obesity related genetic variant (FTO rs9939609) and prostate cancer risk. PLoS One. 2010;5(10):e13485.
24. Lurie G, Gaudet MM, Spurdle AB, et al. The obesity-associated polymorphisms FTO rs9939609 and MC4R rs17782313 and endometrial cancer risk in non-Hispanic white women. PLoS One. 2011;6(2):e16756.
25. Keller L, Xu W, Wang HX, et al. The obesity related gene, FTO, interacts with APOE, and is associated with Alzheimer's disease risk: a prospective cohort study. J Alzheimers Dis. 2011;23(3):461-9.
26. Ho AJ, Stein JL, Hua X, et al. A commonly carried allele of the obesity-related FTO gene is associated with reduced brain volume in the healthy elderly. Proc Natl Acad Sci U S A. 2010 May 4;107(18):8404-9.
27. Bressler J, Fornage M, Demerath EW, et al. Fat mass and obesity gene and cognitive decline: the Atherosclerosis Risk in Communities Study. Neurology. 2013 Jan 1;80(1):92-9.
28. Davis W, van Rensburg SJ, Cronje FJ, et al. The fat mass and obesity-associated FTO rs9939609 polymorphism is associated with elevated homocysteine levels in patients with multiple sclerosis screened for vascular risk factors. Metab Brain Dis. 2014 Jun;29(2):409-19.
29. Belsky DW, Moffitt TE, Sugden K, et al. Development and evaluation of a genetic risk score for obesity. Biodemography Soc Biol. 2013;59(1):85-100.
30. Choh AC, Lee M, Kent JW, et al. Gene-by-age effects on BMI from birth to adulthood: the Fels Longitudinal Study. Obesity (Silver Spring). 2014 Mar;22(3):875-81.
31. Winer S, Paltser G, Chan Y, et al. Obesity predisposes to Th17 bias. Eur J Immunol. 2009 Sep;39(9):2629-35.
32. Brown K, DeCoffe D, Molcan E, et al. Diet-induced dysbiosis of the intestinal microbiota and the effects on immunity and disease. Nutrients. 2012 Aug;4(8):1095-119.
33. Manzel A, Muller DN, Hafler DA, et al. Role of "Western diet" in inflammatory autoimmune diseases. Curr Allergy Asthma Rep. 2014 Jan;14(1):404.
34. Soskic S, Stokic E, Isenovic ER. The relationship between vitamin D and obesity. Curr Med Res Opin. 2014 Jun;30(6):1197-9.
35. Schoindre Y, Terrier B, Kahn JE, et al. [Vitamin D and autoimmunity. First part: Fundamental aspects]. Rev Med Interne. 2012 Feb;33(2):80-6.
36. Cao H. Adipocytokines in obesity and metabolic disease. J Endocrinol. 2014 Feb;220(2):T47-59.

37. Fernandez-Rhodes L, Demerath EW, Cousminer DL, et al. Association of adiposity genetic variants with menarche timing in 92,105 women of European descent. Am J Epidemiol. 2013 Aug 1;178(3):451-60.
38. Kennedy A, Gettys TW, Watson P, et al. The metabolic significance of leptin in humans: gender-based differences in relationship to adiposity, insulin sensitivity, and energy expenditure. J Clin Endocrinol Metab. 1997 Apr;82(4):1293-300.
39. Matarese G, Procaccini C, De Rosa V. The intricate interface between immune and metabolic regulation: a role for leptin in the pathogenesis of multiple sclerosis? J Leukoc Biol. 2008 Oct;84(4):893-9.
40. Mao JH, Langley SA, Huang Y, et al. Identification of genetic factors that modify motor performance and body weight using Collaborative Cross mice. Nature: Sci Rep. 2015 Nov;5:16247.
41. Burgess S, Timpson NJ, Ebrahim S, et al. Mendelian randomization: where are we now and where are we going? Int J Epidemiol. 2015 Apr;44(2):379-88.
42. Barcellos LF, May SL, Ramsay PP, et al. High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. PLoS Genet. 2009 Oct;5(10):e1000696.
43. Li Y, Abecasis GR. Mach 1.0: Rapid Haplotype Reconstruction and Missing Genotype Inference. Am J Hum Genet. 2006 S79:2290.
44. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006 Aug;38(8):904–909.

3.6 Tables and Figures

Table 1. Demographic and disease characteristics of MS cases and controls[a]

Characteristic	KPNC				Sweden			
	MS Cases (N = 1,104)	Controls (N = 10,536)	*P*-value[b]	*P*-value[c]	MS Cases (N = 5,133)	Controls (N = 4,718)	*P*-value[b]	*P*-value[c]
Year of birth	1958 ± 8.93	1958 ± 8.95	0.67	0.72	1960 ± 13.38	1961 ± 13.53	<0.001	0.38
Sex			0.29	0.83			<0.001	0.27
Female	882 (79.89)	8,555 (81.20)			3,741 (73.17)	3,592 (76.13)		
Male	222 (20.11)	1,981 (18.80)			1,392 (27.22)	1,126 (23.87)		
Smoker			<0.001	<0.001			<0.001	0.003
Never	555 (50.36)	7,004 (67.93)			2,163 (42.30)	2,371 (50.25)		
Ever	547 (49.64)	3,307 (32.07)			2,853 (55.80)	2,146 (45.49)		
College graduate			<0.001	0.02			0.001	0.10
Yes	488 (44.20)	3,677 (35.84)			3,725 (72.85)	3,279 (69.50)		
No	616 (55.80)	6,583 (64.16)			1,397 (27.32)	1,433 (30.37)		
*HLA-DRB1*15:01*			<0.001	0.98			<0.001	0.73
0	518 (46.92)	7,752 (73.60)			2,144 (41.93)	3,321 (70.39)		
1-2	586 (53.08)	2,781 (26.40)			2,989 (58.46)	1,397 (29.61)		
wGRS of non-HLA risk variants	12.86 ± 0.68	12.47 ± 0.68	<0.001	0.03	12.48 ± 0.67	12.08 ± 0.69	<0.001	0.05
BMI in young adulthood	22.97 ± 4.37	21.47 ± 3.30	<0.001	<0.001	21.97 ± 3.56	21.68 ± 3.20	<0.001	<0.001
BMI GRS	11.55 ± 0.82	11.47 ± 0.81	0.002	--	11.56 ± 0.82	11.50 ± 0.81	0.0002	--

Percentages may not equal 100 due to missing values.

[a] Table values are mean ± SD for continuous variables and n (column %) for categorical variables.

[b] *P*-value is for t-test (continuous variables) or χ^2 test (categorical variables) between MS Cases vs. Controls for each study (KPNC and Sweden)

[c] *P*-value is for linear regression models for BMI GRS and each variable

Table 2. Association between BMI genetic risk score and self-reported BMI in young adulthood

	KPNC		Sweden	
	Coefficient	R^2 (%)	Coefficient	R^2 (%)
FTO only	0.29 (0.20, 0.39)	0.4	0.27 (0.18, 0.36)	0.3
Females	0.23 (0.13, 0.34)	0.2	0.29 (0.18, 0.40)	0.4
Males	0.51 (0.31, 0.72)	1.2	0.22 (0.06, 0.38)	0.2
BMI GRS	0.64 (0.56, 0.72)	2.3	0.54 (0.45, 0.63)	1.5
Females	0.61 (0.53, 0.70)	2.1	0.51 (0.40, 0.61)	1.3
Males	0.74 (0.57, 0.91)	3.6	0.61 (0.45, 0.76)	2.3

All model *P*-values significant at <0.001

*BMI at age 18 (GERA) or in 20's (KPNC); at age 20 (EIMS/GEMS)

Table 3. Multivariate regression model of BMI on MS susceptibility using BMI GRS*

	KPNC		Sweden	
	OR	***P*-value**	**OR**	***P*-value**
BMI GRS	1.13 (1.04, 1.22)	0.004	1.09 (1.03, 1.15)	0.002
Sex	0.94 (0.80, 1.12)	0.50	0.80 (0.72, 0.89)	<0.001
Year of birth	1.01 (1.00, 1.01)	0.07	1.01 (1.00, 1.01)	0.006
No college education	0.77 (0.67, 0.88)	<0.001	0.91 (0.83, 1.01)	0.07
Ever smoking	2.02 (1.76, 2.31)	<0.001	1.49 (1.36, 1.62)	<0.001
*HLA-DRB1*15:01*	3.55 (3.10, 4.06)	<0.001	3.43 (3.14, 3.76)	<0.001
wGRS of non-HLA MS risk variants	2.32 (2.10, 2.57)	<0.001	2.36 (2.20, 2.52)	<0.001

*Adjusted for ancestry using principal components; Sweden was additionally controlled for study type

Table 4. Significant controlled direct effects of BMI variants on MS susceptibility*

SNP	CHR	GENE	OR (95% CI) KPNC[#]	*P*-value	OR (95% CI) Sweden[#]	*P*-value	OR (95% CI) Meta-analysis
Significant controlled direct effect associated with an increased risk of MS							
rs11126666	2	*KCNK3*	1.20 (0.96, 1.48)	0.002	1.13 (0.93, 1.29)	0.001	1.15 (1.01, 1.30)
rs2112347	5	*POC5*	1.05 (0.83, 1.28)	0.32	1.16 (1.04, 1.30)	<0.001	1.13 (1.02, 1.25)
rs1558902	16	*FTO*	1.13 (0.94, 1.36)	0.01	1.17 (1.02, 1.35)	<0.001	1.16 (1.03, 1.28)
rs7243357	18	*GRP*	1.14 (0.80, 1.49)	0.06	1.26 (1.06, 1.46)	<0.001	1.23 (1.06, 1.40)
Significant controlled direct effect associated with a decreased risk of MS							
rs7599312	2	*ERBB4*	0.77 (0.62, 0.97)	<0.001	0.91 (0.78, 1.09)	0.01	0.85 (0.71, 0.98)

*Mediator defined as BMI during young adulthood, set at 22.0 kg/m^2

[#]Adjusted for smoking, education, year of birth, *HLA-DRB1*15:01*, wGRS of 110 non-HLA MS risk variants, genetic ancestry, and gender; Swedish analysis additionally adjusted for study type. All analyses were bootstrapped with 100 replications. Meta-analysis conducted with bootstrapped estimates, adjusted for random effects.

Table 5. Odds ratios of network sub-scores including established BMI and MS risk loci on MS susceptibility

Gene Networks	OR* (95% CI) KPNC	*P*-value KPNC	OR* (95% CI) Sweden	*P*-value Sweden
TNFRSF14, CDC37, TRAF3, CD40, SLC30A7, LTBR, TNFSF14, RTEL1, TNFSF14	1.10 (1.06, 1.14)	<0.001	1.13 (1.10, 1.15)	<0.001
CD48, ***GNAI3****, BCL10, MALT1, CAMK2G, ADCY3, IL12A, IL2RA, IL12B, IQGAP1,* ***NEXN****, STAT4, TYK2, STAT3, MAPK1, IL7R, IL22RA2, TNGRSF25, CAMK2G, RPS6KA4, CARD11, TNFRSF1A, ERBB2, PIK3E2,* ***TAL1****, NCOA1, TEAD2, ESRRA, CBLB, PHGDH,* ***TYK2, HIP1,*** *TNGRSF1A, REL, CLTC, NFKB1,* ***ERBB4, CLIP1****, RGS14, EPS15L1,* ***TAOK2****, AP1M2^*	1.08 (1.06, 1.10)	<0.001	1.08 (1.07, 1.09)	<0.001
C3, F12	1.17 (1.08, 1.28)	<0.001	1.13 (1.08, 1.18)	<0.001
DMXL2*, RAB3A*	1.18 (1.10, 1.27)	<0.001	1.05 (1.00, 1.10)	0.04
NPEPPS, ***GBE1***	1.07 (1.00, 1.15)	0.04	1.05 (1.01, 1.10)	0.02
MAST3, BAD, ***MAP2K5****,* ***FOXO3, RABEP1***	1.04 (0.99, 1.09)	0.08	1.06 (1.03, 1.09)	<0.001
CD86, CD80	1.06 (0.99, 1.14)	0.09	1.19 (1.14, 1.25)	<0.001
TUFM*, ILF3,* ***LSM4, SNRPD2****,* ***RPL27A****^ +*	1.04 (0.99, 1.09)	0.10	1.02 (0.99, 1.05)	0.13
APOE*, SLC9A8,* ***APOC1, DGKG*** *+*	1.05 (0.98, 1.13)	0.14	1.04 (0.99, 1.08)	0.09
HHEX, SOX8	1.05 (0.97, 1.14)	0.24	1.08 (1.03, 1.13)	0.001

NUP88, POM121C	0.99 (0.93, 1.07)	0.88	1.03 (0.98, 1.07)	0.25
HIC1, TCF7L2	1.00 (0.93, 1.08)	0.97	1.00 (0.96, 1.05)	0.85
NPC1, PDK4	1.00 (0.93, 1.07)	1.00	1.02 (0.98, 1.06)	0.36

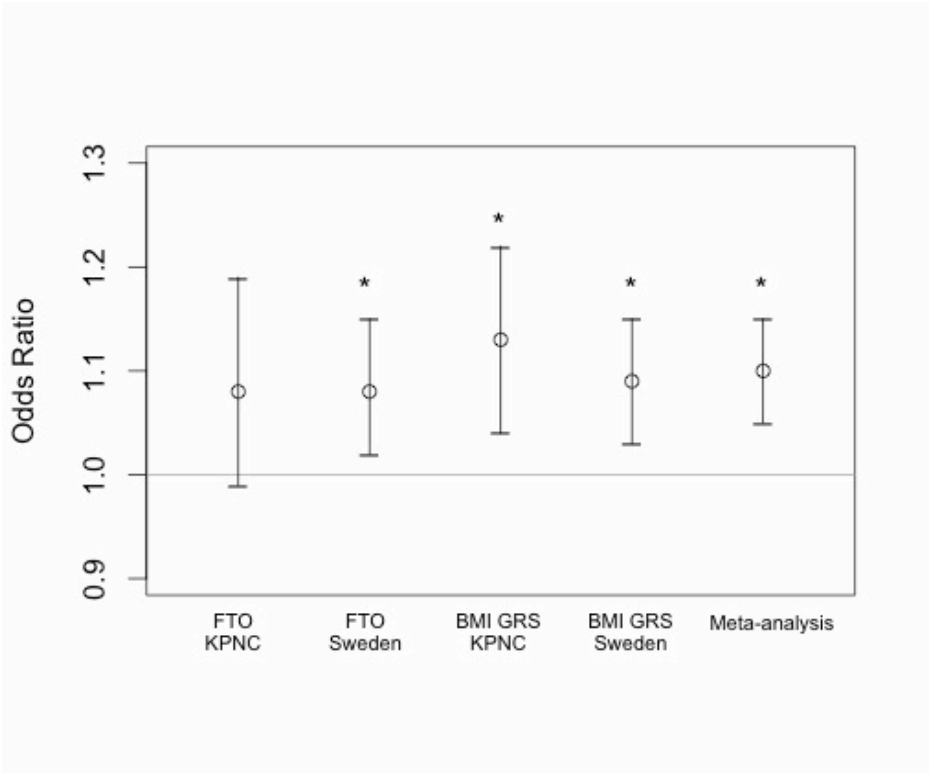
*ORs after controlling for sex, year of birth, genetic ancestry, smoking, and number of *HLA-DRB1*15:01* alleles; Sweden additionally controlled for study type.

N.B. Bolded genes represent those identified from BMI-related variants

^Subscore associated with self-reported BMI during young adulthood in Kaiser (age 18 GERA, 20's KPNC)

+Subscore associated with self-reported BMI during young adulthood in Sweden

Figure 1. Causal odds ratio of BMI on MS susceptibility using genetic variants as instrumental variables



Adjusted for smoking, education, year of birth, *HLA-DRB1*15:01*, wGRS of non-HLA MS risk variants, genetic ancestry, and sex; Sweden was additionally controlled for study type.
**P*-value <0.01

Figure 2. Protein-protein interactions between MS and BMI-related genes

Supplementary Information

KPNC genotyping and quality control

Whole blood was collected, processed and extracted for DNA using Gentra Puregene protocol. Saliva was collected using Oragene kits. Medium resolution *HLA-DRB1* and genome-wide single nucleotide polymorphism (SNP) genotyping was performed as previously described[3,42] using Axiom (Affymetrix) custom chip for GERA controls, and Illumina Infinium 660K BeadChip Array and Human Omni Express for KPNC cases and controls. Low-quality SNPs were removed prior to imputation (<1% minor allele frequency, genotyped in <90% individuals) and samples with >10% failed genotype calls, duplicates or related individuals. Imputation against reference haplotypes from 1,000 Genomes Project was conducted using SHAPEIT and IMPUTE2 (info score >0.8 on all three platforms; minor allele frequency in controls with standard deviation <0.03). Cross-platform association tests were also conducted to remove SNPs associated with the genotype array (false discovery rate q<0.05). Population outliers were identified using multidimensional scaling and reference samples from Human Genome Diversity Project (http://www.hagsc.org/hgdp) and removed from analyses.

Swedish genotyping and quality control

All participants were asked to give blood samples, which were genotyped on an Illumina custom array and on OmniExpress-24. *HLA-DRB1* information was imputed with HLA*IMP02 using genotypes in the major histocompatibility complex region from the custom array. SNPs with <2% minor allele frequency, genotyped in <98% of individuals, or not in Hardy Weinberg equilibrium among controls (P<0.0001) were removed from analysis. Individuals with >2% failed genotype calls, related individuals, or population outliers identified using the SmartPCA program were removed.[43] Twelve BMI SNPs were taken from the custom array, and the remaining genotypes from the OmniExpress chip. Forty-three BMI SNPs were not present on the array and were imputed using MaCH 1.0[44] with standard settings and the Northern European 1,000 Genomes reference panel. Seventeen markers utilized the August 2009 reference panel, 25 markers the August 2010 panel, and 1 marker the July 2011 panel.

Supplementary Table 1. List of genes mapped to BMI and MS risk variants by DAPPLE for network analysis

Gene	SNP
ADCY3	rs4665719
ADPGK	rs7164727
AGBL2	rs3817334
AHI1	rs11154801
ALDOA	rs4787491
AMPD2	rs17024393
ANAPC1	rs17174870
AP1M2	rs2288904
APOC1	rs2075650
APOE	rs2075650
ARFRP1	rs2256814
ATP2A1	rs3888190
ATXN2L	rs3888190
BACH2	rs72928038
BAD	rs694739
BATF	rs4903324
BCKDK	rs9925964
BCL10	rs12087340
BCL9L	rs9736016
BDNF	rs11030104
BOLA2	rs7204270
BOLA2B	rs7204270
C1orf106	rs55838263
C3	rs1077667
C6orf106	rs205262
CADM1	rs12286929
CALCR	rs9641123
CAMK2G	rs2688608
CAPSL	rs6881706
CARD11	rs1843938
CBLB	rs2028597
CCDC155	rs8107548
CCDC88B	rs694739
CCR4	rs4679081

CD40	rs4810485
CD48	rs35967351
CD58	rs6677309
CD5L	rs2050568
CD6	rs34383631
CD80	rs1131265
CD86	rs2255214
CDC37	rs34536443
CDH3	rs1886700
CENPO	rs4665719
CLIP1	rs11057405
CLTC	rs8070345
CLUAP1	rs758747
CNNM2	rs11191560
CNRIP1	rs7595717
CORO1A	rs7204270
CPT1B	rs470119
CTSH	rs59772922
CXCR5	rs9736016
CYP24A1	rs2248359
DGKG	rs1516725
DKKL1	rs8107548
DMXL2	rs3736485
DNAJB4	rs12401738
DNAJC27	rs10182181
DOC2A	rs4787491
DPH5	rs11581062
EAF2	rs1920296
EHBP1	rs11688816
ELAVL4	rs11583200
ELMO1	rs60600003
EPS15L1	rs1870071
ERBB2	rs12946510
ERBB4	rs7599312
ESPN	rs3007421
ESRRA	rs694739
ETV5	rs1516725
ETV7	rs941816

EVI5	rs41286801
F12	rs4976646
FAIM2	rs7138803
FDX1L	rs34536443
FHIT	rs2365389
FNBP4	rs7120737
FOXO3	rs9400239
FOXP1	rs9828629
FUBP1	rs12401738
GALNT10	rs7715256
GBE1	rs3849570
GDPD3	rs7204270
GFI1	rs41286801
GIPR	rs2287019
GLB1	rs4679081
GNAI3	rs17024393
GRB7	rs12946510
GRID1	rs7899106
GRP	rs7243357
GSTM4	rs17024393
HHEX	rs7923837
HHIP	rs11727676
HIC1	rs9914578
HIP1	rs1167827
HIRIP3	rs4787491
HSD17B12	rs2176598
ICAM3	rs34536443
IKZF3	rs12946510
IL12A	rs1014486
IL12B	rs2546890
IL22RA2	rs17066096
IL2RA	rs2104286
IL7R	rs6881706
ILF3	rs2288904
INO80E	rs4787491
IPO9	rs2820292
IQCB1	rs1920296
IQGAP1	rs8042861

IRF8	rs35929052
JAZF1	rs917116
JDP2	rs4903324
KBTBD4	rs3817334
KCNK3	rs11126666
KCTD15	rs29941
KCTD20	rs941816
KIAA2026	rs2150702
LIME1	rs2256814
LMAN2	rs4976646
LMX1B	rs10733682
LRFN2	rs2033529
LSM4	rs17724992
LTBR	rs1800693
MAF	rs7196953
MALT1	rs7238078
MAP2K5	rs16951275
MAPK1	rs2283792
MAPK3	rs7204270
MAST3	rs11554159
MERTK	rs17174870
MIOX	rs470119
MMEL1	rs3748817
MXD3	rs4976646
NAV1	rs2820292
NCAPH2	rs470119
NCOA1	rs4665719
NCOA5	rs4810485
NDUFS3	rs3817334
NEGR1	rs3101336
NEXN	rs12401738
NFKB1	rs7665090
NLRC3	rs758747
NPC1	rs1808579
NPEPPS	rs4794058
NRXN3	rs7141420
NT5C2	rs11191560
NUP88	rs1000940

PARK2	rs13191362
PDK4	rs6465468
PDZK1IP1	rs977747
PFDN4	rs2248359
PFN3	rs4976646
PHGDH	rs666930
PHLDB1	rs533646
PIK3R2	rs11554159
PITPNM2	rs7132277
PLAU	rs2688608
PLEK	rs7595717
PLEKHG5	rs3007421
POM121C	rs1167827
PRDX5	rs694739
PRSS8	rs9925964
PTPRK	rs802734
PTRF	rs4796791
PTRH2	rs8070345
PVRL2	rs2075650
QPCTL	rs2287019
RAB3A	rs11554159
RABEP1	rs1000940
RALYL	rs2033732
RASA2	rs16851483
RASGRF1	rs59772922
RAVER1	rs34536443
RBM17	rs2104286
REL	rs842639
RGS1	rs1359062
RGS14	rs4976646
RIOK3	rs1808579
RPAIN	rs1000940
RPL27A	rs4256980
RPS6KA4	rs694739
RSPH3	rs212405
RTEL1	rs2256814
SBK1	rs2650492
SCARB2	rs17001654

SCNN1A	rs1800693
SCO2	rs470119
SEC16B	rs543874
SH2B1	rs3888190
SKP1	rs756699
SLAMF7	rs35967351
SLC30A7	rs11581062
SLC34A1	rs4976646
SLC9A8	rs17785991
SMG6	rs9914578
SNRPD2	rs2287019
SOX8	rs2744148
SP110	rs9989735
SP140	rs9989735
SPDEF	rs205262
SSTR5	rs2744148
ST5	rs4256980
STAT3	rs4796791
STAT4	rs9967792
STIL	rs977747
TAGAP	rs212405
TAL1	rs977747
TAOK2	rs4787491
TCF7	rs756699
TCF7L2	rs7903146
TEAD2	rs8107548
TET2	rs2726518
TFAP2B	rs2207139
TMEM39A	rs1131265
TNFAIP3	rs67297943
TNFRSF14	rs3748817
TNFRSF1A	rs1800693
TNFRSF25	rs3007421
TNFSF14	rs1077667
TOMM40	rs2075650
TRAF3	rs12148050
TREH	rs533646
TUFM	rs3888190

TYK2	rs34536443
UBAC2	rs4772201
UBLCP1	rs2546890
USP37	rs492400
VCAM1	rs7552544
VIL1	rs492400
VPS33A	rs11057405
WWOX	rs12149527
ZFP36L1	rs2236262
ZGPAT	rs2256814
ZMIZ1	rs1782645
ZNF646	rs9925964

Supplementary Table 2. Direct connections of genes identified through network analysis

Gene 1 -- Gene 2	Gene 1 Type	Gene 2 Type
AP1M2 -- CLTC	MS	MS
APOE -- SLC9A8	BMI	MS
BCL10 -- MALT1	MS	MS
C3 -- F12	MS	MS
CAMK2G -- ADCY3	MS	MS
CARD11 -- BCL10	MS	MS
CARD11 -- MALT1	MS	MS
CBLB -- IL22RA2	MS	MS
CBLB -- IL2RA	MS	MS
CBLB -- IL7R	MS	MS
CBLB -- NEXN	MS	BMI
CBLB -- PHGDH	MS	MS
CBLB -- TYK2	MS	MS
CD48 -- GNAI3	MS	BMI
CD86 -- CD80	MS	MS
CLIP1 -- IQGAP1	BMI	MS
CLTC -- HIP1	MS	BMI
CLTC -- MAPK1	MS	MS
CLTC -- MAPK3	MS	MS
CLTC -- REL	MS	MS
CLTC -- TNFRSF1A	MS	MS
DMXL2 -- RAB3A	BMI	MS
EPS15L1 -- CLTC	MS	MS
ERBB2 -- MAPK1	MS	MS
ERBB2 -- STAT3	MS	MS
ERBB4 -- ERBB2	BMI	MS
ERBB4 -- PIK3R2	BMI	MS
GNAI3 -- ADCY3	BMI	MS
GNAI3 -- MAPK1	BMI	MS
HHEX -- SOX8	MS	MS
HIC1 -- TCF7L2	BMI	BMI
IL12A -- IL12B	MS	MS
IL12A -- IL2RA	MS	MS
IL12B -- IL2RA	MS	MS
IL22RA2 -- IL12A	MS	MS
IL22RA2 -- IL12B	MS	MS
IL22RA2 -- TYK2	MS	MS
IL2RA -- STAT3	MS	MS
IL2RA -- TYK2	MS	MS

IL7R -- IL12A	MS	MS
IL7R -- IL12B	MS	MS
IL7R -- STAT3	MS	MS
IL7R -- TYK2	MS	MS
ILF3 -- SNRPD2	MS	BMI
IQGAP1 -- NEXN	MS	BMI
LSM4 -- SNRPD2	BMI	BMI
LTBR -- TNFSF14	MS	MS
LTBR -- TRAF3	MS	MS
MAP2K5 -- BAD	BMI	MS
MAPK1 -- CAMK2G	MS	MS
MAPK1 -- IQGAP1	MS	MS
MAPK1 -- STAT3	MS	MS
MAPK1 -- TNFRSF25	MS	MS
MAPK3 -- ERBB2	MS	MS
MAPK3 -- MAPK1	MS	MS
MAPK3 -- RPS6KA4	MS	MS
MAPK3 -- STAT3	MS	MS
MAST3 -- BAD	MS	MS
MAST3 -- FOXO3	MS	BMI
MAST3 -- RABEP1	MS	BMI
NCOA1 -- ESRRA	MS	MS
NCOA1 -- MAPK1	MS	MS
NCOA1 -- STAT3	MS	MS
NCOA1 -- TEAD2	MS	MS
NFKB1 -- CLTC	MS	MS
NFKB1 -- IL12A	MS	MS
NFKB1 -- IL12B	MS	MS
NFKB1 -- IL2RA	MS	MS
NFKB1 -- NCOA1	MS	MS
NFKB1 -- PIK3R2	MS	MS
NFKB1 -- REL	MS	MS
NFKB1 -- STAT3	MS	MS
NPC1 -- PDK4	BMI	BMI
NPEPPS -- GBE1	MS	BMI
NUP88 -- POM121C	BMI	BMI
PIK3R2 -- CARD11	MS	MS
PIK3R2 -- ERBB2	MS	MS
PIK3R2 -- TYK2	MS	MS
RGS14 -- GNAI3	MS	BMI
RGS14 -- IQGAP1	MS	MS
RPL27A -- ILF3	BMI	MS
RPS6KA4 -- MAPK1	MS	MS

SLC30A7 -- CD40	MS	MS
SLC9A8 -- APOC1	MS	BMI
SLC9A8 -- DGKG	MS	BMI
STAT4 -- TYK2	MS	MS
TAL1 -- MAPK3	BMI	MS
TAOK2 -- NEXN	BMI	BMI
TNFRSF14 -- CDC37	MS	MS
TNFRSF1A -- BCL10	MS	MS
TNFRSF1A -- MAPK1	MS	MS
TNFRSF1A -- TNFRSF25	MS	MS
TNFSF14 -- RTEL1	MS	MS
TNFSF14 -- TNFRSF14	MS	MS
TRAF3 -- CD40	MS	MS
TRAF3 -- TNFRSF14	MS	MS
TRAF3 -- TNFSF14	MS	MS
TUFM -- ILF3	BMI	MS
TYK2 -- STAT3	MS	MS

Chapter 4

Vitamin D, body mass index, and pediatric-onset multiple sclerosis: evidence for a causal independent association

4.1 Background

Multiple sclerosis (MS) is a severe autoimmune inflammatory disease of the central nervous system and affects approximately 400,000 individuals in the United States.[1] Although disease onset typically occurs between the ages of 20 and 40, approximately 5% of all MS patients have symptom onset before 18 years[2-4]. Exact mechanisms involved in pediatric MS pathogenesis are unknown; however, similar to adult MS, infection with Epstein-Barr virus, exposure to cigarette smoking, and the genetic risk factor *HLA-DRB1* are associated with pediatric-onset MS.[2, 5, 6] While there is strong research demonstrating the association between low vitamin D status and increased risk of adult-onset MS,[7, 8] there is a paucity of studies examining this relationship in pediatric MS. Additionally, an association between childhood obesity and risk of pediatric MS has been reported,[9] though a causal relationship remains to be confirmed.

Mendelian randomization is a type of instrumental variable (IV) analysis that uses genetic variants strongly associated with an exposure, rather than a direct measure of the exposure, to estimate the effect of the exposure on an outcome. Because inherited alleles are not affected by most potential confounding variables or by disease status, the typical confounding present in observational studies does not hinder IV analyses and reverse causation is unlikely. Our goal was to estimate the causal association of vitamin D and BMI on pediatric-onset MS risk using IV analysis based on genetic variants from established large genome-wide association studies (GWAS) in two large datasets of cases and controls.

4.2 Methods

U.S. Participants

Pediatric-onset MS cases (n= 738) originated from three cohorts. Most were enrolled through pediatric MS centers established at UC San Francisco, Stony Brook, Children's Hospital of Philadelphia, Texas Children's Hospital, University of Colorado School of Medicine, University of Texas Southwestern, State University of New York Buffalo, Loma Linda University, Mayo Clinic, University of Alabama at Birmingham, Ann & Robert Lurie Children's Hospital of Chicago, University of Utah, Boston Children's Hospital, Brigham and Women's Hospital, Washington University St. Louis, and Children's National Medical Center between January 2006 and December 2014. These centers are tertiary referral centers, but also serve regional patients from all socioeconomic groups. Consecutive patients with onset of MS or clinically isolated syndrome (CIS) suggestive of early MS before age 18 seen at some of these pediatric MS clinics were offered participation as previously described.[10] In addition, we utilized adult cases from Kaiser Permanente Northern California with reported age of onset < 18 years. The third cohort of distinct cases were enrolled in a NIH supported case-control study (NS071463, PI Waubant) with MS or CIS onset before the age of 18, seen within 4 years of disease onset, with at least 2 silent

T2-bright foci on brain and cord MRI, and ascertained by a panel of at least two pediatric MS experts.[11] Some cases were also provided from another MS genetic study and only patients with a disease onset prior to age 18 based on medical records were included (Dr Oksenberg). Established diagnostic criteria were used for all cases.[12, 13]

Because pediatric-onset MS is a rare condition, control individuals were derived from multiple sources to increase statistical power. These included: (1) adult controls (self-report White, non-Hispanic) recruited from the KPNC membership without a diagnosis of MS or related condition (optic neuritis, transverse myelitis, or demyelination disease) confirmed through electronic medical records (N=1,182);[14] (2) adult controls from the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort who participated in the KPNC Research Program on Genes, Environment, and Health (RPGEH) without evidence of MS in EMR (dbGaP phs000674.v2.p2; N= 10,819); (3) pediatric controls (R01 NS071463, PI Waubant; N=70); and (4) pediatric controls (multiple race/ethnicities) recruited as part of the Northern California Childhood Leukemia Study (N=229).

All study protocols were approved by the Institutional Review Boards for Human Subjects at UC San Francisco, Stony Brook, Children's Hospital of Philadelphia, Texas Children's Hospital, University of Colorado School of Medicine, University of Texas Southwestern, State University of New York Buffalo, Loma Linda University, Mayo Clinic, University of Alabama at Birmingham, Ann & Robert Lurie Children's Hospital of Chicago, University of Utah, Boston Children's Hospital, Brigham and Women's Hospital, Washington University St. Louis, Children's National Medical Center, Kaiser Permanente Division of Research and UC Berkeley. Informed consent or assent (children) was obtained for all study participants and their parents when appropriate.

Whole blood was collected, processed and extracted for DNA using Gentra Puregene protocol or received as Guthrie Card blood samples and extracted for DNA using the QIAamp DNA Micro Kit for Dried Blood Spots. Saliva was collected for DNA extraction using Oragene kits.

Genome wide profiling was performed for all pediatric-onset MS cases and a subset of controls using the Illumina Infinium platform. The Infinium 660K BeadChip or HumanOmniExpressExom BeadChip was used to genotype each study participant. An additional control group (RPGEH) was genotyped using custom designed Affymetrix Axiom arrays.

Classical multidimensional scaling (MDS) was performed to visualize population substructure and provide quantitative measures (components) of population genetic variation. MDS was completed for all genotyped data, which was merged with the Human Genome Diversity Project (HGDP) reference. SNPs present in <90% were removed. Analysis of genome-wide average proportion of alleles shared identical by state (IBS) was performed using PLINK v.1.07, and related/identical individuals were removed. Study samples were aligned with HGDP references, and the first three dimensions from MDS of the HGDP European population was used. Individual outliers were identified using MDS and excluded.

Characterization of *HLA-DRB1*15* (0, 1 or 2 allele copies) for each study participant was based on the rs3135388 tag SNP (allele). The SNP is highly correlated with *DRB1*15* alleles

(r^2=0.97).[15] Twenty-four case samples were dropped due to onset ≥ 18 or duplicate samples. We restricted our sample to white non-Hispanics, the largest group in our dataset, to ensure a genetically homogenous sample and avoid the possibility of confounding by ancestry. After excluding population outliers, the final dataset was comprised of 394 pediatric-onset MS cases and 10,875 controls.

Swedish Participants

Data were collected from two population based case-control studies on incident (Epidemiological Investigation of Risk Factors for MS [EIMS]) and prevalent (Genes and Environment in MS [GEMS]) MS patients with reported age of onset < 18 years. The EIMS study (2005-2014) inclusion criteria were: age 16-70 years, diagnosed MS according to the to the McDonald criteria (15, 16) within two years, and ability to understand the Swedish language. GEMS study participants were identified from the Swedish National MS registry, fulfilled the McDonald criteria (15, 16), and were recruited during 2009-2011. For both studies, controls were randomly chosen from the population register and matched to cases by sex, age at inclusion in the study, and region of residence. Two controls were matched to each case in the EIMS study and one control per case in the GEMS study. All participants in the EIMS study were distinct from those in the GEMS study. Ethical approval for both studies was obtained from the Regional Ethical Review Board in North Stockholm and participants provided written informed consent at Karolinska Institutet. Details of the study design have been described elsewhere (7, 12). The participation rate in the EIMS study was 92% for cases and 67% for controls, and in the GEMS study 82% for cases and 66% for controls. Genotyping data were available for 75% of EIMS and 91% of GEMS participants. Data for 275 cases and 5,376 controls was available from the Swedish studies.

Swedish genotyping and quality control

All participants were asked to give blood samples, which were genotyped on an Illumina custom array and on OmniExpress-24. *HLA-DRB1* information was imputed with HLA*IMP02 using genotypes in the major histocompatibility complex region from the custom array. SNPs with <2% minor allele frequency, genotyped in <98% of individuals, or not in Hardy Weinberg equilibrium among controls (P<0.0001) were removed from analysis. Individuals with >2% failed genotype calls, related individuals, or population outliers identified using the SmartPCA program were removed (41). Ninety of the 110 MS risk SNPs were present on the custom array. Proxy SNPs for an additional eighteen MS risk SNPs were selected from the custom array (R^2 >0.8, except RS28723576, which had R^2 >0.71). All three Vitamin D SNPs were present on the custom array. Twelve BMI SNPs were taken from the custom array, and the remaining genotypes from the OmniExpress chip. Forty-three BMI SNPs were not present on the array and were imputed using MaCH 1.0 (42) with standard settings and the Northern European 1,000 Genomes reference panel. Seventeen markers utilized the August 2009 reference panel, 25 markers the August 2010 panel, and 1 marker the July 2011 panel. The variant rs2245368 could not be imputed with good quality and was therefore omitted from the BMI GRS.

Exposure assessment

A previous GWAS identified three SNPs strongly associated with serum vitamin D level: rs2282679, rs2060793, and rs3829251, which together explain approximately 2.8% of the variance in circulating vitamin D levels.[16] A genetic IV for vitamin D (vitD GRS) was

constructed using the three risk variants, weighting each allele by the effect size seen in the GWAS and summing across the variants. The BMI IV was derived using beta coefficients as weights for 97 variants associated with BMI (R^2=2.7%) identified through the largest and most recent GWAS for BMI.[17] The sum of risk alleles multiplied by the estimated effect reported of each risk allele on the phenotype was calculated as each individual's BMI genetic risk score (BMI GRS).

We also explored whether variants specifically associated with childhood BMI (chBMI) were associated with pediatric-onset MS susceptibility. A chBMI GRS was constructed using 28 independent variants associated with pediatric/childhood BMI in the literature.[18-24] The score was unweighted, and calculated by summing the number of risk alleles across each loci. Of the 28 variants in the chBMI GRS, 11 overlapped or were highly correlated with ($R^2 > 0.6$) variants in the BMI GRS.

A weighted genetic risk score (wGRS) that combines the weighted odds ratio (OR) from each of 110 non-MHC MS susceptibility loci identified through recent GWAS[25] was calculated for each pediatric-onset case and control by multiplying the number of risk alleles for each loci by the weight for that variant and then taking the sum across the 110 loci.[26] One SNP was missing for the US study (rs201202118), and two SNPs were missing for Sweden (rs2028597, rs6874308). The weight for each locus is the natural log of the odds ratio for each allele.

Statistical Analyses

Logistic regression was used to estimate the effect of each GRS on pediatric-onset MS case/control status, controlling for sex, wGRS, presence of any *HLA-DRB1*15:01* alleles, and ancestry as determined by MDS. 95% confidence intervals (CI), odds ratio (OR) and *P*-values were reported. Interaction was assessed on the multiplicative scale. Meta-analysis assuming fixed effects was performed if chi-squared tests of heterogeneity demonstrated $P > 0.05$. All analyses were conducted in PLINK, STATA (StataCorp, College Station, TX), or R.

4.3 Results

Demographic features and other clinical characteristics of cases and controls are shown in Table 1. Mean age at onset for pediatric cases was 14.05 years (+/- 3.30) in the US study and 16.0 (+/- 2.33) in the Swedish study. There was a significant difference in wGRS and *HLA-DRB1*15:01* status between cases and controls in both studies (P <0.001). Cases had a lower vitD GRS and a compared to controls, though this difference was not significant in the US study (P_{US}=0.11 and P_{Sweden}=0.03). Cases on average carried had a higher BMI GRS than controls, but the difference was only significant in the US population (P_{US}=0.002 and P_{Sweden}=0.49); there was no significant difference in chBMI GRS between cases and controls in either study population. There was also no association between age of onset and any of the IVs in either study population (data not shown).

In the meta-analysis of both US and Swedish studies, vitD GRS was significantly associated with a decreased risk of MS (OR=0.75, 95% CI 0.59, 0.97, P=0.03) after adjusting for sex, *HLA-DRB1*15:01*, wGRS, and genetic ancestry derived from principal components (Table 2). The IV analysis for the BMI GRS also demonstrated a causal association between BMI and pediatric-

onset MS as represented by the GRS of 97 BMI variants after adjusting for covariates (OR = 1.15, 95% CI 1.04, 1.27; *P*=0.01). There was no significant association between chBMI GRS and pediatric-onset MS (OR=1.01, 95% CI 0.97, 1.04, *P*=0.30).

There was no evidence of interaction between vitD GRS or chBMI GRS and *HLA-DRB1*15:01* (data not shown); however, a significant interaction was present between BMI GRS and *HLA-DRB1*15:01* in the US study (*P*-interaction = 0.04). Individuals carrying 1-2 *DRB1*15:01* risk alleles demonstrated a stronger association (OR = 1.39) compared to non-carriers (OR=1.05). However, this interaction was not present in the Swedish dataset (*P*-interaction=0.66).

When modeled together, both the vitD and BMI GRS appeared to independently contribute to pediatric-onset MS susceptibility. Joint estimates were consistent with those found when each GRS was modeled alone in both populations (data not shown).

4.4 Discussion

For the first time, we report a *causal* and independent association between low serum concentrations of vitamin D and increased BMI, and risk of pediatric-onset MS after adjusting for sex, ancestry, *HLA-DRB1*15:01,* and 110 non-HLA MS risk variants. While previous studies have attributed an increased risk of MS associated with BMI to lower vitamin D levels seen in obese individuals, our findings show that these risk factors may in fact *independently* contribute to pediatric-onset disease susceptibility.

Several studies have shown a protective effect of vitamin D levels on MS risk. Two prospective studies showed a significantly reduced risk of MS in those with high 25(OH)D.[27, 28] Maternal vitamin D deficiency (25[OH]D levels < 12.02 ng/mL) during early pregnancy may be associated with a two-fold risk of MS in offspring.[29] Further, a recent MR study showed a causal effect for low 25(OH)D on adult MS risk.[30] Findings from our meta-analysis reveal for the first time that vitamin D as represented as a GRS are significantly associated with pediatric-onset MS. We believe that a larger sample size or stronger genetic instrument with additional variants may help to demonstrate an even stronger effect, as each SNP used in the vitD GRS explained $\leq$ 1.2% of the variance in serum 25(OH)D concentrations.

A large body of evidence suggests that high BMI in both childhood and adolescence is associated with adult MS risk.[14, 31, 32] Our meta-analysis results reflect this finding, and suggest that increased BMI is a causal risk factor for pediatric MS disease onset. Additionally, childhood obesity and risk of pediatric[9] onset MS has been reported. We did not find any association between the chBMI GRS and pediatric-onset MS; however, this may be due to weak instrument bias. Interestingly, research has also indicated that HLA genes interact with self-reported BMI during adolescence to increase the risk of MS,[33] an interaction that was also present in our study based on a BMI GRS. Stratification based on *HLA-DRB1*15:01* indicated a significant interaction with BMI GRS in the US study, with increased risk for those carrying 1-2 vs. 0 risk alleles, though we were unable to replicate this finding in the Swedish dataset.

There are several hypotheses linking low vitamin D and high BMI to autoimmune diseases, including MS. Increased exposure to vitamin D leads to decreased production of inflammatory

cytokines, a decrease in Th1 and Th17 cell differentiation, and an increase in T regulatory cells, suggesting that low vitamin D is acting on MS by shifting the balance of the immune response toward a more pro-inflammatory state.[34-36] Obesity may also induce a Th17 response via an IL-6 dependent process leading to exacerbation of inflammatory diseases such as MS.[37] Additionally, a Th17/Treg imbalance may lead to alteration of intestinal microbiome in obese individuals, altering the intestinal immune response which may explain the association between increased BMI and MS.[38, 39]

Our study had several strengths including clinically well-characterized pediatric-onset MS cases with an onset about 20 years earlier than most adult MS cases used in large genetic studies, rigorous diagnostic criteria and focus on a single racial group. Comprehensive methods were used in the current study to identify population outliers and remove them prior to analysis. We also analyzed BMI and vitamin D risk scores together, and used two study samples to conduct a large, well powered meta-analysis. By utilizing genetic risk scores as a proxy for exposure, reverse causation and confounding by unmeasured or unknown factors are unlikely.

Limitations include identification of pediatric-onset MS cases through tertiary clinics, which may not represent all MS cases in children. It is possible that individuals with a more benign disease or without access to healthcare may have been missed. Additionally, IV analysis involves many assumptions that we have addressed as best as possible. We were able to meet most model assumptions by utilizing a vitD and BMI GRS constructed from weights derived from independent populations through large GWAS. However, associations of genetic variants with unmeasured or unknown confounders cannot be ruled out. Our study included non-Hispanic whites, which may limit the generalizability of our findings. Additional limitations include a small male sample size, assumption of linearity, and possible pleiotropic effects of vitD and BMI GRS on MS.

We provide evidence for the first time that low vitamin D and increased BMI may truly be causally and independently associated with pediatric-onset MS. The effect of low vitamin D and obesity may involve independent predisposing genetic factors and biological pathways mediating disease onset that future studies will unravel.

4.5 References

1. Hogancamp WE, Rodriguez M, Weinshenker BG. The epidemiology of multiple sclerosis. Mayo Clin Proc 1997;72:871-878.
2. Chitnis T, Krupp L, Yeh A, et al. Pediatric multiple sclerosis. Neurol Clin 2011;29:481-505.
3. Yeh EA, Chitnis T, Krupp L, et al. Pediatric multiple sclerosis. Nature reviews Neurology 2009;5:621-631.
4. Krupp LB, Tardieu M, Amato MP, et al. International Pediatric Multiple Sclerosis Study Group criteria for pediatric multiple sclerosis and immune-mediated central nervous system demyelinating disorders: revisions to the 2007 definitions. Multiple sclerosis 2013.
5. Mikaeloff Y, Caridade G, Tardieu M, Suissa S. Parental smoking at home and the risk of childhood-onset multiple sclerosis in children. Brain 2007;130:2589-2595.
6. Waldman A, Ghezzi A, Bar-Or A, Mikaeloff Y, Tardieu M, Banwell B. Multiple sclerosis in children: an update on clinical diagnosis, therapeutic strategies, and research. Lancet Neurol 2014;13:936-948.
7. Banwell B, Bar-Or A, Arnold DL, et al. Clinical, environmental, and genetic determinants of multiple sclerosis in children with acute demyelination: a prospective national cohort study. Lancet Neurol 2011;10:436-445.
8. Mowry EM, Krupp LB, Milazzo M, et al. Vitamin D status is associated with relapse rate in pediatric-onset multiple sclerosis. Ann Neurol 2010;67:618-624.
9. Langer-Gould A, Brara SM, Beaber BE, Koebnick C. Childhood obesity and risk of pediatric multiple sclerosis and clinically isolated syndrome. Neurology 2013;80:548-552.
10. Waubant E, Mowry EM, Krupp L, et al. Common viruses associated with lower pediatric multiple sclerosis risk. Neurology 2011;76:1989-1995.
11. McDonald J, Graves J, Waldman A, et al. A case-control study of dietary salt intake in pediatric-onset multiple sclerosis. Mult Scler Relat Disord 2016;6:87-92.
12. Krupp LB, Tardieu M, Amato MP, et al. International Pediatric Multiple Sclerosis Study Group criteria for pediatric multiple sclerosis and immune-mediated central nervous system demyelinating disorders: revisions to the 2007 definitions. Mult Scler 2013;19:1261-1267.
13. McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol 2001;50:121-127.
14. Gianfrancesco MA, Acuna B, Shen L, et al. Obesity during childhood and adolescence increases susceptibility to multiple sclerosis after accounting for established genetic and environmental risk factors. Obes Res Clin Pract 2014;8:e435-447.
15. Zivkovic M, Stankovic A, Dincic E, et al. The tag SNP for HLA-DRB1*1501, rs3135388, is significantly associated with multiple sclerosis susceptibility: cost-effective high-throughput detection by real-time PCR. Clin Chim Acta 2009;406:27-30.
16. Ahn J, Yu K, Stolzenberg-Solomon R, et al. Genome-wide association study of circulating vitamin D levels. Hum Mol Genet 2010;19:2739-2745.
17. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197-206.
18. Bradfield JP, Taal HR, Timpson NJ, et al. A genome-wide association meta-analysis identifies new childhood obesity loci. Nat Genet 2012;44:526-531.

19. den Hoed M, Ekelund U, Brage S, et al. Genetic susceptibility to obesity and related traits in childhood and adolescence: influence of loci identified by genome-wide association studies. Diabetes 2010;59:2980-2988.
20. Manco M, Dallapiccola B. Genetics of pediatric obesity. Pediatrics 2012;130:123-133.
21. Mitchell JA, Hakonarson H, Rebbeck TR, Grant SF. Obesity-susceptibility loci and the tails of the pediatric BMI distribution. Obesity (Silver Spring) 2013;21:1256-1260.
22. Scuteri A, Sanna S, Chen WM, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet 2007;3:e115.
23. Stergiakouli E, Gaillard R, Tavare JM, et al. Genome-wide association study of height-adjusted BMI in childhood identifies functional variant in ADCY3. Obesity (Silver Spring) 2014;22:2252-2259.
24. Warrington NM, Howe LD, Paternoster L, et al. A genome-wide association study of body mass index across early life and childhood. Int J Epidemiol 2015;44:700-712.
25. Beecham AH, Patsopoulos NA, Xifara DK, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 2013;45:1353-1360.
26. De Jager PL, Chibnik LB, Cui J, et al. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurol 2009;8:1111-1119.
27. Munger KL, Levin LI, Hollis BW, Howard NS, Ascherio A. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. Jama 2006;296:2832-2838.
28. Salzer J, Hallmans G, Nystrom M, Stenlund H, Wadell G, Sundstrom P. Vitamin D as a protective factor in multiple sclerosis. Neurology 2012;79:2140-2145.
29. Munger KL, Aivo J, Hongell K, Soilu-Hanninen M, Surcel HM, Ascherio A. Vitamin D Status During Pregnancy and Risk of Multiple Sclerosis in Offspring of Women in the Finnish Maternity Cohort. JAMA Neurol 2016;73:515-519.
30. Mokry LE, Ross S, Ahmad OS, et al. Vitamin D and Risk of Multiple Sclerosis: A Mendelian Randomization Study. PLoS Med 2015;12:e1001866.
31. Hedstrom 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 2012;18:1334-1336.
32. Munger KL, Chitnis T, Ascherio A. Body size and risk of MS in two cohorts of US women. Neurology 2009;73:1543-1550.
33. Hedstrom AK, Lima Bomfim I, Barcellos L, et al. Interaction between adolescent obesity and HLA risk genes in the etiology of multiple sclerosis. Neurology 2014;82:865-872.
34. Aranow C. Vitamin D and the immune system. J Investig Med 2011;59:881-886.
35. Hart PH, Gorman S, Finlay-Jones JJ. Modulation of the immune system by UV radiation: more than just the effects of vitamin D? Nat Rev Immunol 2011;11:584-596.
36. Prietl B, Treiber G, Pieber TR, Amrein K. Vitamin D and immune function. Nutrients 2013;5:2502-2521.
37. Winer S, Paltser G, Chan Y, et al. Obesity predisposes to Th17 bias. Eur J Immunol 2009;39:2629-2635.
38. Brown K, DeCoffe D, Molcan E, Gibson DL. Diet-induced dysbiosis of the intestinal microbiota and the effects on immunity and disease. Nutrients 2012;4:1095-1119.
39. Manzel A, Muller DN, Hafler DA, Erdman SE, Linker RA, Kleinewietfeld M. Role of "Western diet" in inflammatory autoimmune diseases. Curr Allergy Asthma Rep 2014;14:404.

Table 1. Demographic and disease characteristics of pediatric MS cases and controls

	US			Sweden		
	Pediatric MS Cases (N = 394)	Controls (N = 10,875)	*P*-value	Pediatric MS Cases (N = 262)	Controls (N = 5,376)	*P*-value
Age of Onset	14.05 ± 3.30	--	--	16.0 ± 2.33	--	--
Females:Males	3.0:1	1.6:1	0.02	2.8:1	3.2:1	0.36
wGRS	12.78 ± 0.67	12.31 ± 0.69	<0.001	13.04 ± 0.68	12.51 ± 0.70	<0.001
*HLA-DRB1*15:01*			<0.001			<0.001
0	193 (49)	7,973 (72)		94 (36)	3,806 (71)	
1-2	201 (51)	2,902 (28)		168 (64)	1,570 (29)	
vitD GRS	1.02 ± 0.28	1.04 ± 0.31	0.11	1.03 ± 0.33	1.06 ± 0.32	0.03
BMI GRS	11.51 ± 0.85	11.38 ± 0.81	0.002	11.54 ± 0.81	11.50 ± 0.81	0.49
chBMI GRS	25.09 ± 3.10	24.80 ± 3.23	0.08	25.83 ± 3.49	25.91 ± 3.33	0.72

Table values are mean ± SD for continuous variables and n (column %) for categorical variables; *P*-value is for t-test or χ^2 test between cases and controls

Table 2. Instrumental variable meta-analysis results demonstrating a causal association between genetic risk scores and pediatric MS susceptibility

	Odds Ratios (95% CI)	*P*-value
Vitamin D GRS	0.75 (0.59, 0.97)	0.03
BMI GRS	1.15 (1.04, 1.27)	0.01
Childhood BMI GRS	1.01 (0.99, 1.04)	0.30

Analyses adjusted for sex, *HLA-DRB1*15:01*, wGRS, and genetic ancestry derived from principal components

*Fixed effects meta-analysis; all chi-squared tests of heterogeneity $P > 0.05$

Chapter 5

Two adipokines demonstrate independent pathways associated with multiple sclerosis susceptibility

5.1 Introduction

MS is an immune-mediated, demyelinating disorder with widespread axonal degeneration occurring throughout the disease process. Both genetic and environmental factors contribute to disease susceptibility.[1] Recent studies have identified several MS genetic risk factors including the *HLA-DRB1*15:01* allele within the major histocompatibility complex (MHC),[2, 3] and 110 variants outside of the MHC in individuals of European ancestry[3]. Established environmental risk factors associated with MS risk include exposure to tobacco smoke, Epstein-Barr Virus infection, and low levels of vitamin D.[4]

In addition, the relationship between obesity and MS risk has been well established,[5] including evidence of a causal association[6] and interaction with HLA genes.[7] However, the biological mechanism underlying this relationship is still unknown. Several links between obesity and autoimmunity have been studied and received recent attention,[8] including the actions of adipokines, or cytokines released from adipose tissue, which are involved in immunity and inflammation. It has been found that white adipose tissue produces over 50 different adipokines, though only a small number have been implicated in autoimmunity.

Research within the MS field has primarily focused on the role of leptin, a key pro-inflammatory adipokine, on disease susceptibility in experimental autoimmune encephalitis (EAE), an animal model of MS. In one study, leptin was required to promote EAE and contribute to disease progression.[9] In human studies, MS cases were found to have higher serum and cerebrospinal fluid levels of leptin compared to controls,[10] and leptin mRNA were shown to be upregulated in active MS lesions.[11] Soluble leptin receptor (sOBr) levels reflect main leptin-binding activity in human blood, and are thought to act as a potential reservoir of bioactive leptin. Additionally, expression of sOBr has been found to be higher in CD8+T cells and monocytes from relapsing-remitting MS patients in a relapse compared to patients in remission and in controls.[12] However, results from these studies may be a reflection of the disease manifestation and indicate reverse causation. The manner in which leptin and sOBr may contribute to disease susceptibility in MS is still unknown.

Unlike leptin, the role of adiponectin in autoimmune disease is less clear. This adipokine has demonstrated both pro- and anti-inflammatory effects. High adiponectin is inversely correlated with risk of type 2 diabetes, coronary artery disease, stroke, and several metabolic traits.[13] Adiponectin-deficient EAE also exhibit worse clinical and histological disease compared to wild-type mice.[14] Calorie restriction in EAE has been shown to be correlated with a decrease in leptin concentrations and IL-6, increased adiponectin levels, and improved outcomes.[15] However, high serum concentrations of adiponectin have been found to be elevated in humans with various autoimmune inflammatory conditions,[16, 17] and there has been some evidence of high concentrations in MS cases.[18]

Lastly, the role of resistin in autoimmune disease is not well defined and has not been extensively studied in MS. One study found no association between a polymorphism in a resistin gene promoter (rs1862513) and risk of MS.[19] However, effects of MS treatments on adipokine levels found that decreased resistin levels correlated with immunomodulatory treatment in relapsing remitting patients,[20] and higher levels are observed in primary progressive MS patients compared to other subgroups, as well as controls.[21]

While animal models of adipokines are informative and shed light on potential biological mechanisms contributing to disease susceptibility and severity, human studies assessing exposure of these immunological markers on MS may be more informative to treatment interventions. Using genetic variants as a proxy of exposure allows one to estimate the effect of various adipokines on MS susceptibility in a manner that isn't confounded by disease status, as the exposure is assigned at birth. No study has previously examined the role of genetic variants associated with several adipokines on MS status in a comprehensive manner. We investigated the relationship between three adipokines (sOBr, adiponectin and resistin) and MS susceptibility using genome-wide significant variants associated with serum concentration levels of these cytokines in a large dataset of cases and controls.

5.2 Methods

Participants
MS cases and controls were recruited from members of Kaiser Permanente Medical Care Plan, Northern California Region (KPNC). KPNC is an integrated health services delivery system with a membership of 3.2 million that comprises about 25-30% of the population of a 22 county service area and is the largest healthcare provider in northern California. Membership is objectively representative of the general population; however, persons in impoverished neighborhoods are underrepresented.[22]

This study was restricted to self-identified white (non-Hispanic) race/ethnicity, the population with the highest prevalence of MS. Eligible KPNC cases were defined as: individuals with a diagnosis of MS by a neurologist (ICD9 code 340.xx; 94.7% of cases had at least two MS diagnoses by a neurologist), age of 18 through 69 years, and membership in KPNC at initial contact. A total of 3,293 potential MS cases were reviewed by KPNC neurologists, who approved contact with 2,823 (86%) at the time of the data freeze (August 2014). Diagnoses were validated utilizing electronic health record (EHR) review and published diagnostic criteria.[23, 24]

Controls were KPNC current members without a diagnosis of MS or related condition (optic neuritis, transverse myelitis, or demyelination disease) confirmed through electronic records, and white (non-Hispanic) race/ethnicity. Controls were matched to cases on age, gender, and zip code. Potential study participants were contacted by mail with a follow-up phone call to explain the study and procedures. The participation rate was approximately 80% for cases and 66% for controls. Genetic data were available for approximately 80% of study participants.

Additional controls included individuals in the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort who participated in the KPNC Research Program on Genes, Environment, and Health (RPGEH). The RPGEH was established to research genetic and

environmental influences on health and disease and is described in more detail elsewhere (dbGaP phs000674.v2.p2). GERA cohort members completed a broad written consent and provided a saliva sample for DNA extraction. Approximately 77% returned completed consent forms for placement in dbGaP, resulting in a final sample size of 78,486 participants. From these participants, we selected a subset of 12,605 self-reported non-Hispanic white individuals without evidence of MS in EHR matched to cases on gender and age (+/- 2 years) at a 10:1 ratio. Study protocols were approved by the Institutional Review Boards (IRB) of KPNC and the University of California, Berkeley.

Exposure Assessment

KPNC study participants completed a computer-assisted telephone interview (CATI) administered by trained staff interviewers and comprised of questions related to various events and exposures as described elsewhere.[25] GERA controls completed a survey consisting of questions related to health behaviors, sociodemographic information, and diagnoses (dbGaP phs00674.v1.p1). KPNC study participants reported their highest and lowest (non-pregnancy) weight during their 20's. Mean weight of each KPNC participant during their 20's was calculated by averaging the highest and lowest weight reported. GERA controls reported their weight at age 18. Each participant's body mass index (BMI) was calculated by dividing mean weight in kilograms by height reported at time-of-interview in meters squared.

Genotyping and quality control

Whole blood was collected, processed and extracted for DNA using Gentra Puregene protocol. Saliva was collected using Oragene kits. Medium resolution *HLA-DRB1* and genome-wide single nucleotide polymorphism (SNP) genotyping was performed as previously described[2, 26] using Axiom (Affymetrix) custom chip for GERA controls, and Illumina Infinium 660K BeadChip Array and Human Omni Express for KPNC cases and controls. Low-quality SNPs were removed prior to imputation (<1% minor allele frequency, genotyped in <90% individuals) and samples with >10% failed genotype calls, duplicates or related individuals. Imputation against reference haplotypes from 1,000 Genomes Project was conducted using SHAPEIT and IMPUTE2 (info score >0.8 on all three platforms; minor allele frequency in controls with standard deviation <0.03). Cross-platform association tests were also conducted to remove SNPs associated with the genotype array (false discovery rate $q<0.05$). Population outliers were identified using multidimensional scaling and reference samples from Human Genome Diversity Project (http://www.hagsc.org/hgdp) and removed from analyses.

We identified three adipokines to measure their relationship with MS susceptibility (sOBr, adiponectin, and resistin) based on recent literature citing their potential importance in autoimmune disease.[8] Genome-wide significant variants associated with serum levels of these three adipokines were collected: sOBr (4 SNPs),[27] adiponectin (12 SNPs),[13] and resistin (5 SNPs)[28] (Supplementary Table 1). One sOBr variant, rs1751492 was in strong linkage disequilibrium, or correlated with, rs1137100 and thus was excluded from analyses. All other variants were independent from one another ($R^2 < 0.6$).

Statistical Analyses

A weighted genetic risk score (wGRS) of MS risk variants was calculated for each individual that weights risk alleles by the logarithm of the odds ratio for each of the 110 non-HLA MS

susceptibility loci identified through the most recent MS GWAS.[3] The wGRS was calculated by multiplying the number of risk alleles for each locus by the weight for that variant, and then taking the sum across the 110 loci. One SNP was missing from the score analysis (rs201202118).

We additionally controlled for BMI and vitamin D serum levels using instrumental variables (IV) to assess independent effects of the adipokine variants on MS susceptibility after accounting for these factors. A genetic IV for vitamin D (vitD GRS) was constructed using the three risk variants (rs2282679, rs2060793, and rs3829251), weighting each allele by the effect size seen in the GWAS[29] and summing across the variants. The BMI IV was derived using beta coefficients as weights for 97 variants associated with BMI identified through the largest and most recent GWAS for BMI.[30] The sum of risk alleles multiplied by the estimated effect reported of each risk allele on the phenotype was calculated as each individual's BMI genetic risk score (BMI GRS).

After quality control and removal of population outliers, a total of 2,162 individuals from KPNC (1,103 cases, 804 controls) with genetic data were available as well as an additional 9,728 controls from the GERA study for a total of 1,103 cases and 10,532 controls. Demographic differences between cases and controls were compared using X^2 tests and independent sample t-tests where appropriate. Logistic regression was used to demonstrate the association between each adipokine SNP and MS status. All analyses were controlled for sex, year of birth, ever smoking, college education, *HLA-DRB1*15:01*, wGRS of non-HLA MS risk variants, and ancestry as derived from MDS components. Bonferroni adjustment was use to corrected for multiple testing using. We also additionally calculated interaction on the multiplicative scale to assess whether significant adipokines interacted with sex and *HLA-DRB1*15:01*. Odds ratios with 95% CIs were estimated. Analyses were conducted in PLINK or STATA. This study was focused on a single hypothesis established a priori; therefore, we report 95% confidence intervals and use an α=0.05 threshold for statistical significance.

5.3 Results

Demographic and disease characteristics of MS cases and controls are found in Table 1. There were significant differences between cases and controls with respect to smoking, college graduation, and *HLA-DRB1*15:01* status, and wGRS (P<0.001). BMI during young adulthood was also significantly higher for MS cases than controls (22.97 $\pm$ 4.37 vs. 21.47 $\pm$ 3.30, respectively; P<0.001).

Results demonstrated a significant association for three sOBr SNPs in the *LEPR* gene and MS susceptibility, including the strongest odds ratio (OR) for rs2767485 (OR=1.26, 95% CI 1.12, 1.42; P<0.0001) and missense variant rs1137100 (OR=1.16, 95% CI 1.04, 1.29; P=0.01) (Table 2). Additionally, two adiponectin SNPs (rs2925979 and rs998584) were significantly associated with MS after adjusting for covariates (OR= 1.15, 95% CI 1.03, 1.28; P=0.01 and OR = 1.19, 95% CI 1.06, 1.34; P=0.003, respectively). An association between one resistin SNP (rs6068258) and MS was also found (OR = 1.14, 95% CI 1.03, 1.27; P=0.01). All associations remained significant after adjustment for BMI and vitamin D levels using IVs, suggesting independent effects associated with MS that do not act through these pathways (P<0.05).

We found no strong evidence of correlation between the vitamin D and BMI instruments, and any of the adipokine variants in controls (Pearson's r < |0.04|). Two SNPs, rs2767485 in *LEPR* and rs998584 in *VEGFA*, remained significant after adjustment for multiple testing (P<0.05). Stratification by sex, as well as *DRB1*15:01* status, did not show any evidence of interaction with any of the adipokine variants (Supplementary Table 2). Multivariate modeling with both rs2767485 and rs998584 together also demonstrated significant independent effects of sOBr and adiponectin on MS susceptibility after controlling for other established risk factors (Table 3).

5.4 Discussion

This study was the first to comprehensively examine the relationship between genetic variants associated with adipokines and MS susceptibility in a large case-control dataset. Results suggest that significant variants hold after adjustment for BMI and vitamin D, and may indicate that the actions of adipokines, specifically sOBr and adiponectin, may have a direct effect on MS susceptibility independent of these factors. Because genetic variants are assigned at exposure and do not suffer from reverse causation, our results suggest a causal association between levels of adipokines and MS susceptibility.

Only one variant in *LEPR* associated with decreased sBOr levels remained significant after multiple testing correction (rs2767485). Our results demonstrate a role for this adipokine in MS susceptibility, and are in line with previous findings. sOBr can modulate steady-state leptin levels after binding free leptin in circulation, and thus lower sBOr levels are associated with higher circulating leptin levels. Studies in mouse models of MS have shown that high levels of leptin are associated with disease onset and progression[9] and correlate with development of pathogenic T cell responses.[31] Findings are also in agreement with observational studies of MS that measured differences in leptin levels between cases and controls,[10] and may indicate a causal role for low levels of sOBr on disease susceptibility in MS. Although sBOr levels are typically inversely correlated with BMI,[32] we found that rs2767485 significantly increased the risk of MS independent of BMI status. Previous studies have also shown that leptin levels are higher in females than males,[33] suggesting a role for leptin in explaining the female predominance of autoimmune diseases such as MS. While we did not see any evidence of interaction between sex and *LEPR* variants, we cannot rule out other potential environmental influences on leptin levels that may increase the risk of MS; for example, adipokines such as leptin are up-regulated by ovarian sex steroids. While one study found no significant difference in genetic variation of three SNPs in the *LEPR* gene (rs1137100, rs1137101, and rs8179183) between MS cases and controls,[34] the sample size was much smaller than the one utilized in our analysis.

In addition, our results showed a significant association for increasing serum adiponectin levels and increased odds of MS specifically associated with rs998584 in *VEGFA*. This variant has previously been associated with increased triglyceride and decreased HDL levels.[35] However, our results are in contrast with studies a studied that showed decreased adiponectin levels in EAE models.[14] Our study was the first to examine the role of genetic variants associated with adiponectin levels on MS onset, a measure less confounded by disease activity or progression. Findings suggest that the association between adiponectin and MS disease may be complex, and warrant further investigation. Future studies should consider prospective measures of adiponectin in order to better determine whether there is a causal effect of this adipokine on risk of MS.

A major strength of this study is the large sample size. We also conducted rigorous quality control to account for population stratification, and we were able to model both genetic and environmental risk factors associated with disease susceptibility. Additional strengths include the fact that cases and controls were sampled from the same source population (i.e. Kaiser Permanente, Northern California patient population). Our study was the first to examine the role several adipokines on MS susceptibility in a comprehensive manner while controlling for multiple established risk factors in the context of sophisticated causal modeling.

This study included non-Hispanic whites, which limits the generalizability of our findings. Additional limitations of our study include a relatively small male sample size and possible pleiotropic effects of genes related to adipokines on MS (i.e. genes may influence phenotypes other than adipokine levels that are associated with an increased risk of MS). The genetic variants in our study explain only a small percentage of the variance in adipokine levels. For example, the five sOB-R variants jointly explained 4.6% of the total variation of sOB-R levels,[27] with rs2767485 explaining approximately 2% after adjusting for other SNPs and covariates, such as fasting status, age, BMI, and diabetes status. Additional variants and environmental factors may influence sOB-R and other adipokine levels, and should be explored in the future. Future studies should also aim to replicate findings, specifically in populations of other race/ethnicities, and better examine how specific adipokine-related variants may influence MS susceptibility and severity.

In conclusion, we conducted the largest comprehensive assessment of adipokine genetic variants and MS susceptibility. Our results are consistent with previous studies demonstrating an association between certain adipokines and MS in human and animal models. By utilizing genetic variants associated with serum levels of adipokines, we were able to measure unconfounded associations between these cytokines and MS. Results demonstrate that adipokines influence disease onset independently of BMI and vitamin D, and offer future directions for treatment mechanisms in MS susceptibility.

5.5 References

1. Favorova OO, Kulakova OG and Boiko AN. [Multiple sclerosis as a polygenic disease: an update]. *Genetika*. 2010; 46: 302-13.
2. Barcellos LF, Sawcer S, Ramsay PP, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. *Human molecular genetics*. 2006; 15: 2813-24.
3. Beecham AH, Patsopoulos NA, Xifara DK, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet*. 2013; 45: 1353-60.
4. Ascherio A and Munger K. Epidemiology of multiple sclerosis: from risk factors to prevention. *Seminars in neurology*. 2008; 28: 17-28.
5. Munger KL, Chitnis T and Ascherio A. Body size and risk of MS in two cohorts of US women. *Neurology*. 2009; 73: 1543-50.
6. Gianfrancesco M, Glymour M, Walter S, et al. Genetic variants associated with body mass index demonstrate a causal effect on multiple sclerosis susceptibility. *Am J Epidemiol*. 2016; In press.
7. Hedstrom AK, Lima Bomfim I, Barcellos L, et al. Interaction between adolescent obesity and HLA risk genes in the etiology of multiple sclerosis. *Neurology*. 2014; 82: 865-72.
8. Versini M, Jeandel PY, Rosenthal E and Shoenfeld Y. Obesity in autoimmune diseases: not a passive bystander. *Autoimmun Rev*. 2014; 13: 981-1000.
9. Matarese G, Di Giacomo A, Sanna V, et al. Requirement for leptin in the induction and progression of autoimmune encephalomyelitis. *J Immunol*. 2001; 166: 5909-16.
10. Matarese G, Carrieri PB, La Cava A, et al. Leptin increase in multiple sclerosis associates with reduced number of CD4(+)CD25+ regulatory T cells. *Proc Natl Acad Sci U S A*. 2005; 102: 5150-5.
11. Lock C, Hermans G, Pedotti R, et al. Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nat Med*. 2002; 8: 500-8.
12. Frisullo G, Mirabella M, Angelucci F, et al. The effect of disease activity on leptin, leptin receptor and suppressor of cytokine signalling-3 expression in relapsing-remitting multiple sclerosis. *J Neuroimmunol*. 2007; 192: 174-83.
13. Dastani Z, Hivert MF, Timpson N, et al. Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet*. 2012; 8: e1002607.
14. Piccio L, Cantoni C, Henderson JG, et al. Lack of adiponectin leads to increased lymphocyte activation and increased disease severity in a mouse model of multiple sclerosis. *Eur J Immunol*. 2013; 43: 2089-100.
15. Piccio L, Stark JL and Cross AH. Chronic calorie restriction attenuates experimental autoimmune encephalomyelitis. *J Leukoc Biol*. 2008; 84: 940-8.
16. Otero M, Lago R, Gomez R, et al. Changes in plasma levels of fat-derived hormones adiponectin, leptin, resistin and visfatin in patients with rheumatoid arthritis. *Ann Rheum Dis*. 2006; 65: 1198-201.
17. Sada KE, Yamasaki Y, Maruyama M, et al. Altered levels of adipocytokines in association with insulin resistance in patients with systemic lupus erythematosus. *J Rheumatol*. 2006; 33: 1545-52.
18. Hietaharju A, Kuusisto H, Nieminen R, Vuolteenaho K, Elovaara I and Moilanen E. Elevated cerebrospinal fluid adiponectin and adipsin levels in patients with multiple sclerosis: a Finnish co-twin study. *Eur J Neurol*. 2010; 17: 332-4.

19. Hossein-Nezhad A, Varzaneh FN, Mirzaei K, Emamgholipour S and Sahraian MA. A polymorphism in the resistin gene promoter and the risk of multiple sclerosis. *Minerva Med*. 2013; 104: 431-8.
20. Michalak S, Jernas L, Wysocka E, et al. The effect of methylprednisolone, interferon beta and glatiramer acetate treatment on the levels of leptin, adiponectin, and resistin in multiple sclerosis. *Neurology and Neurophysiology*. 2014: S12.
21. Emamgholipour S, Eshaghi SM, Hossein-nezhad A, Mirzaei K, Maghbooli Z and Sahraian MA. Adipocytokine profile, cytokine levels and foxp3 expression in multiple sclerosis: a possible link to susceptibility and clinical course of disease. *PLoS One*. 2013; 8: e76555.
22. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health*. 1992; 82: 703-10.
23. McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol*. 2001; 50: 121-7.
24. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of neurology*. 2011; 69: 292-302.
25. Gianfrancesco MA, Acuna B, Shen L, et al. Obesity during childhood and adolescence increases susceptibility to multiple sclerosis after accounting for established genetic and environmental risk factors. *Obes Res Clin Pract*. 2014; 8: e435-47.
26. Barcellos LF, May SL, Ramsay PP, et al. High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. *PLoS Genet*. 2009; 5: e1000696.
27. Sun Q, Cornelis MC, Kraft P, et al. Genome-wide association study identifies polymorphisms in LEPR as determinants of plasma soluble leptin receptor levels. *Hum Mol Genet*. 2010; 19: 1846-55.
28. Qi Q, Menzaghi C, Smith S, et al. Genome-wide association analysis identifies TYW3/CRYZ and NDST4 loci associated with circulating resistin levels. *Hum Mol Genet*. 2012; 21: 4774-80.
29. Ahn J, Yu K, Stolzenberg-Solomon R, et al. Genome-wide association study of circulating vitamin D levels. *Hum Mol Genet*. 2010; 19: 2739-45.
30. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015; 518: 197-206.
31. Sanna V, Di Giacomo A, La Cava A, et al. Leptin surge precedes onset of autoimmune encephalomyelitis and correlates with development of pathogenic T cell responses. *J Clin Invest*. 2003; 111: 241-50.
32. van Dielen FM, van't Veer C, Schols AM, Soeters PB, Buurman WA and Greve JW. Increased leptin concentrations correlate with increased concentrations of inflammatory markers in morbidly obese individuals. *Int J Obes Relat Metab Disord*. 2001; 25: 1759-66.
33. Kennedy A, Gettys TW, Watson P, et al. The metabolic significance of leptin in humans: gender-based differences in relationship to adiposity, insulin sensitivity, and energy expenditure. *J Clin Endocrinol Metab*. 1997; 82: 1293-300.
34. Rey LK, Wieczorek S, Akkad DA, Linker RA, Chan A and Hoffjan S. Polymorphisms in genes encoding leptin, ghrelin and their receptors in German multiple sclerosis patients. *Mol Cell Probes*. 2011; 25: 255-9.
35. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013; 45: 1274-83.

Table 1. Demographic and disease characteristics of MS cases and controls[a]

Characteristic	MS Cases (N = 1,103)	Controls (N = 10,532)	*P*-value[b]
Year of birth	1958 ± 8.94	1958 ± 8.95	0.67
Sex			0.29
Female	881 (80)	8,553 (81)	
Male	222 (20)	1,979 (19)	
Age at first symptom	31.87 ± 9.73	--	--
MS Subtype		--	--
Relapsing Remitting	674 (66)		
Secondary Progressive	161 (16)		
Primary Progressive	84 (8)		
Relapsing Progressive	32 (3)		
Don't Know	66 (6)		
Smoker			<0.001
Never	554 (50)	7,001 (70)	
Ever	547 (50)	3,306 (30)	
College graduate			<0.001
Yes	488 (44)	3,677 (36)	
No	615 (56)	6,579 (64)	
*HLA-DRB1*15:01*			<0.001
0	517 (47)	7,750 (74)	
1-2	586 (53)	2,779 (26)	
wGRS of 110 non-HLA risk variants	12.86 ± 0.68	12.47 ± 0.68	<0.001
BMI at age 18 (GERA) or in 20's (KPNC)			<0.001
<18.5	89 (8)	1,257 (13)	
18.5-<21	308 (29)	3,639 (39)	
21-<23	248 (23)	2,272 (24)	
23-<25	169 (16)	1,185 (13)	
25-<27	95 (9)	553 (6)	
27-<30	71 (7)	320 (3)	
≥30	77 (7)	222 (1)	

[a] Table values are mean ± SD for continuous variables and n (column %) for categorical variables. Percentages may not equal 100 due to rounding.

[b] *P*-value is for t-test (continuous variables) or χ^2 test (categorical variables) between MS Cases vs. Controls

Table 2. Results of multivariate regression models of significant individual adipokine variants on MS susceptibility

			Adjusted[a]		Adjusted[a] + BMI GRS + Vit D GRS	
Gene	SNP	Effect Allele	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Soluble leptin receptor levels						
LEPR	rs1137100	A	1.16 (1.03, 1.29)	0.01	1.16 (1.04, 1.29)	0.01
LEPR	rs1137101	A	1.11 (1.01, 1.22)	0.04	1.11 (1.01, 1.22)	0.03
LEPR	rs2767485	C	1.26 (1.12, 1.42)	<0.0001[b]	1.26 (1.12, 1.42)	<0.0001[b]
Adiponectin						
CMIP	rs2925979	T	1.14 (1.03, 1.27)	0.01	1.15 (1.03, 1.28)	0.01
VEGFA	rs998584	C	1.20 (1.07, 1.35)	0.003[b]	1.20 (1.07, 1.35)	0.003[b]
Resistin						
LINC01524	rs6068258	G	0.88 (0.79, 0.97)	0.01	0.88 (0.80, 0.98)	0.02

[a]Adjusted for *HLA-DRB1*15:01*, wGRS, sex, and genetic ancestry, education, smoking, year of birth

[b]Significant after multiple testing correction

Table 3. Adjusted estimates demonstrating independent associations between significant adipokines and MS susceptibility

	Covariates*		Covariates, BMI GRS, and Vitamin D GRS*	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
rs2767485, *LEPR* (soluble leptin receptor level)	1.26 (1.12, 1.41)	0.0002	1.26 (1.12, 1.42)	0.0001
rs998584, *VEGFA* (adiponectin)	1.20 (1.06, 1.35)	0.003	1.20 (1.06, 1.35)	0.003
Smoker (0=no, 1=yes)	2.11 (1.84, 2.42)	<0.0001	2.09 (1.83, 2.40)	<0.0001
*HLA-DRB1*15:01* positive (0=no, 1=yes)	3.71 (3.24, 4.25)	<0.0001	3.72 (3.25, 4.26)	<0.0001
wGRS	2.34 (2.11, 2.58)	<0.0001	2.33 (2.11, 2.58)	<0.0001
BMI GRS	--	--	1.14 (1.05, 1.24)	0.001
Vitamin D GRS	--	--	0.83 (0.67, 1.03)	0.09

*Additionally adjusted for sex, year of birth, education and genetic ancestry

Supplementary Table 1. Genetic variants associated with adipokines through genome-wide association studies

Chr	Gene	SNP	Effect Allele	Type
Variants associated with decreased soluble leptin receptor levels				
1	*LEPR*	rs1137100	A	Missense
		rs1137101	A	Missense
		rs2767485	C	Intron
		rs1751492*	T	Intron
		rs4655555	T	Intron
Variants associated with increased adiponectin levels				
1	*LYPLAL1*	rs3001032	C	Upstream gene variant
3	*GNL3*	rs1108842	C	5 prime UTR variant
3	*ADIPOQ*	rs6810075	T	Downstream gene variant
3	*TSC22D2*	rs1597466	G	Intergenic
6	*VEGFA*	rs998584	C	Downstream gene variant
8	*TRIB1*	rs2980879	T	Intron
12	*GPR109A*	rs601339	G	Intron variant
12	*DNAH10*	rs7133378	A	Intron
12	*PDE3A*	rs7955516	C	Intergenic
16	*CDH13*	rs12922394	C	Intron
16	*CMIP*	rs2925979	T	Intron
19	*PEPD*	rs731839	A	Intron
Variants associated with increased resistin levels				
1	*C1orf168*	rs17372114	T	Intron
1	*TYW3*	rs3931020	C	Downstream gene variant
4	*NDST4*	rs13144478	T	Intergenic
19	*RETN*	rs3745367	A	Intron
20	*LINC01524*	rs6068258	G	Intron

* In LD with rs1137100

Supplementary Table 2. Results of stratified multivariate regression model of significant individual adipokine variants on MS susceptibility based on sex and *HLA-DRB1* status

		Stratified by Sex			Stratified by *HLA-DRB1* Status		
Gene	SNP	Male OR (95% CI)	Female OR (95% CI)	*P*-value Interaction[#]	0 alleles (95% CI)	1-2 alleles (95% CI)	*P*-value Interaction[#]
Soluble leptin receptor levels							
LEPR	rs1137100	1.18 (0.91, 1.52)	1.16 (1.02, 1.31)	0.68	1.07 (0.92, 1.26)	1.23 (1.05, 1.43)	0.36
LEPR	rs1137101	1.13 (0.91, 1.40)	1.10 (0.99, 1.22)	0.72	1.08 (0.94, 1.24)	1.12 (0.98, 1.28)	0.74
LEPR	rs2767485	1.27 (0.96, 1.67)	1.27 (1.12, 1.45)	0.85	1.23 (1.04, 1.45)	1.31 (1.11, 1.55)	0.68
Adiponectin							
CMIP	rs2925979	1.08 (0.84, 1.38)	1.17 (1.04, 1.32)	0.44	1.16 (0.99, 1.35)	1.15 (0.99, 1.33)	0.76
VEGFA	rs998584	1.17 (.0.89, 1.53)	1.20 (1.05, 1.37)	0.69	1.24 (1.05, 1.48)	1.14 (0.96, 1.34)	0.85
Resistin							
LINC01524	rs6068258	0.97 (0.76, 1.23)	0.85 (0.76, 0.96)	0.38	0.90 (0.78, 1.05)	0.84 (0.72, 0.97)	0.48

*Adjusted for wGRS, genetic ancestry, education, smoking, year of birth and sex or *HLA-DRB1*15:01* (when not stratified).

[#] Interaction assessed on the multiplicative scale

Chapter 6

Conclusion

Multiple sclerosis (MS) is a severe and complex demyelinating disorder resulting in significant disability and decreased quality of life. Strong evidence supports the contribution of both genetic and environmental factors to disease susceptibility. Recently, obesity has emerged as a risk factor for MS; however, the biological mechanism through which obesity and MS may be related is unknown. This dissertation aimed to replicate the association between childhood and adolescent obesity and MS onset, identify a causal relationship, and examine whether specific biological pathways underlie this association using data from both pediatric and adult-onset cases and a large number of population-based controls.

6.1 Summary of Findings

Findings of Chapter 2: Obesity during childhood and adolescence increases susceptibility to MS after accounting for established genetic and environmental risk factors.

This chapter examined the relationship between increased BMI and MS susceptibility, but for the first time additionally controlled for genetic factors, including 110 non-MHC MS risk variants. Results demonstrated a two-fold increased risk of MS in females with a BMI $\geq$ 30 kg/m^2 in one's twenties. No significant association was found in males. Findings also showed a null association between baseline BMI (assessed at interview) and MS, consistent with previously literature. Additionally, a significant association was shown for body size at age 10 (little/very overweight vs. just about right) and MS in females, indicating that both childhood and adolescent obesity contribute to increased susceptibility to MS.

Findings of Chapter 3: Genetic variants associated with body mass index demonstrate a causal effect on MS susceptibility.

Results from this chapter found the first evidence of a causal association between increased BMI and MS susceptibility using an instrumental variable composed of 97 genetic variants associated with BMI. Findings were confirmed in a replication dataset of MS cases and controls from Sweden. In addition, a meta-analysis of the two studies showed that five variants demonstrated evidence of a direct effect on MS susceptibility independent of BMI, indicating that other pathways may mediate disease onset. Results also revealed significant evidence for direct and indirect protein-protein interactions. Five of thirteen interactions were significantly associated with MS susceptibility, and three of these included both obesity and MS genes.

Findings of Chapter 4: Vitamin D, body mass index, and pediatric MS risk: evidence for a causal independent association.

Chapter 4 estimated the causal association between low serum vitamin D concentrations and increased BMI on pediatric-onset MS using instrumental variable analyses. Meta-analysis findings of two populations, from the US and Sweden, demonstrated that a vitamin D genetic

risk score associated with increasing levels of 25[OH]D in serum decreased the odds of pediatric-onset MS after controlling for sex, genetic ancestry, *HLA-DRB1*15:01* and 110 non-HLA MS risk variants. A significant association between BMI GRS and pediatric disease onset was also demonstrated after adjusting for covariates. Estimates were unchanged when both genetic risk scores were modeled together. For the first time, I provide evidence supporting decreased vitamin D levels and increased BMI are *causally* and *independently* associated with pediatric-onset MS.

Findings of Chapter 5: Specific adipokines may underlie the association between obesity and MS susceptibility.

In this chapter, I identified genome-wide significant variants associated with serum levels of three adipokines to measure their relationship with MS: plasma soluble leptin receptor, adiponectin, and resistin. Results demonstrated a significant association for three plasma soluble leptin receptor variants, two adiponectin variants, and one resistin variant after controlling for covariates. The associations remained significant after adjustment for BMI and vitamin D using instrumental variables. Two variants, rs2767485 in *LEPR* and rs998584 in *VEGFA*, remained significant after multiple testing adjustment. No evidence of significant interaction between any of the adipokines and sex or *HLA-DRB1*15:01* was present. These results indicate that MS risk may involve predisposing genetic factors for adipokines independent of BMI and vitamin D, and suggest that independent biological mechanisms may mediate disease onset.

6.2 Conclusions and Future Directions

This dissertation focused on establishing that increased BMI is a causal risk factor for both pediatric and adult-onset MS, with potential evidence of direct effects of BMI genes on MS susceptibility in adults. Findings also indicated that vitamin D and BMI causally increase risk of pediatric-onset MS via independent pathways; further research is needed to understand how these biological pathways operate to induce disease susceptibility. Lastly, genes associated with cytokines released by fat tissues, such as leptin, adiponectin and resistin, also may be associated with disease susceptibility through pathways not mediated by BMI. Thus, self-reported BMI may not entirely capture the mechanisms through which obesity may influence MS susceptibility. Further investigation of the of biological mechanisms of BMI genes identified via bioinformatics databases and/or animal models of MS will inform pathways for new treatment interventions in MS. Taken together, these results indicate that reducing the obesity burden is critically important to reduce MS disease risk. Given that obesity is modifiable, interventions targeted at reducing the obesity burden in the United States and around the world have large public health implications.