LONDON’S GLOBAL UNIVERSITY

UCL

Using Genetic Approach to Investigate the Causal Relationship between Environmental Risk Factors and Multiple Sclerosis

By
Mona Mohammad Almramhi

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London

UCL Queen Square Institute of Neurology

May 2022

Declaration of Authorship

I, Mona Mohammad Almramhi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Over the past few years, the study of genetics has been revolutionised by rapid progress in genome-wide association studies (GWAS), leading to the uncovering of a large proportion of genetic variants associated with biomarkers, lifestyle factors and disease incidence. However, understanding how these variants mechanistically influence disease phenotypes and/or translating GWAS findings into drug targets have proved challenging.

The challenges facing epidemiological studies in distinguishing between causation and association have drawn much interest to the Mendelian randomisation (MR) approach. MR can be viewed as a platform to integrate novel genetic information generated in GWAS or molecular studies (QTL) to inform about causal associations between an exposure and a disease outcome; the demonstrated causal association is less likely to be affected by confounding or reverse causation.

This advancement led me to apply the MR approach to uncover the causal link between several environmentally modifiable exposures and the risk of developing multiple sclerosis (MS), as well as the severity of MS. Additionally, I broadened our approach by applying MR to the druggable genome to identify and prioritise new drug targets for MS.

Overall, I found genetic evidence that high-density lipoprotein and a range of features linked to obesity (body mass index, weight, fat mass and fat percentage) and stroke, are risk factors for MS development. Additionally, I found genetic evidence supports the casual role of obesity in worsening MS severity. Most importantly, this thesis prioritises several genes (*RAC2, CCR4, SLAMF7 and SIK3*) with the potential to serve as druggable genes in MS. This finding offers a platform for informing the design of MS preventive strategies.

Impact Statement

Multiple sclerosis (MS) is a multifactorial disease influenced by both genetic and environmental factors. MS is the most common neurological disability, and about 2.8 million people live with MS worldwide. Thus far, MS has no cure, and the primary goal of licenced disease-modifying therapies (DMTs) is to modify the disease's course and slow disability. However, DMTs are hampered by potentially serious adverse reactions and require careful monitoring through a specialist MS clinic. In addition, many MS DMTs cost beyond US $90,000 annually, which is a major deterrent, particularly in low- and middle-income countries. Presently, a need exists for more research to discover new and better MS interventions and preventive measures. Nevertheless, discovering novel drugs for human diseases is a lengthy, complex and costly process.

Although the aetiology of MS remains unclear, observational studies have reported associations between several risk factors and increased MS risk; however, whether these associations are causal remains unclear.

Accurate assessment of environmental risk factors is crucial not only to understand MS risk aetiology but also to improve prevention strategies, identify novel therapeutic targets and delay disease progression.

By leveraging the availability of large-scale GWASs by applying the MR approach, this thesis contributes to uncovering causal relationships between several environmentally modifiable exposures and the risk of developing MS and the risk of worsening MS severity, which has received less attention in the research field. Interestingly, this thesis sheds light on the possible critical role of ischemic stroke in increasing the risk of developing MS.

Demonstrating causation would aid the understanding of these factors' roles in MS pathogenesis, enrich our MS aetiology knowledge and may carry meaningful implications for MS patient care by aiding clinical diagnosis—and perhaps treatment.

Drug repurposing is an alternative strategy with the potential to complement traditional drug discovery by mitigating high monetary- and time-related costs. By performing MR of the druggable genome, this thesis also provides valuable evidence to identify opportunities for MS prevention and informs on potential drug targets.

This thesis is novel in its application of MR methods to demonstrate how human genetic variation can provide insights into potential risk factors and inform drug design to improve MS patient health.

Acknowledgements

First, I want to express significant gratitude to my supervisor, Professor Nicholas Wood (my principal supervisor), for allowing me to undertake this interdisciplinary PhD with no epidemiology experience and for believing I could perform the work—in doing so, you have introduced me to epidemiology, a field I now thoroughly enjoy. I appreciate Professor Nicholas Wood's boundless enthusiasm, support, advice and encouragement.

This thesis would not have been possible without many different people who have helped me throughout the years. I offer sincere appreciation to Catherine Storm, with whom I have had the pleasure of working and becoming close professional colleagues and friends.

To Dr. Chris Finan, I am thankful that my first steps into this PhD were with you. It is a true privilege to learn from you. To Dr. Floriaan Schmidt, thank you for your invaluable guidance on MR analysis. To Dr. Doug Speed, thank you for your invaluable statistical guidance.

Additionally, I am grateful to my colleagues, Demis Kia, Rachel Coneys, Burleen Chhatwal and Sandesh Chopade, for their support.

I also, would like to express my deepest appreciation to faculty of Applied Medical Sciences at King Abdulaziz University (Jeddah, Saudi Arabia) for awarding me this scholarship to support of my graduate studies. I also extend my warmest and most sincere thanks to the Saudi Arabian Cultural Bureau in London for their assistance and support.

Finally, this thesis is dedicated with my warmest love and thanks to my parents, my husband (Dr. Abdulhamid Fatani), my sister, my brothers, and my little son Tamim, who all believed in me and are extremely proud of me and my achievements. I would not have succeeded without your permeant encouragement to be braver.

Publications

Almramhi M, Storm CS, Kia DA ,Coneys R, Chhatwal B,Wood NW. "The role of body fat in multiple sclerosis susceptibility and severity: A Mendelian randomisation study". *Multiple Sclerosis Journal* [accepted].2022.

Almramhi MM, Finan C, Storm C, Schmidt AF, Kia DA, Coneys R, Chopade S, Hingorani AD, Nicholas NW. Exploring the Role of Plasma Lipids and Statins Interventions on Multiple Sclerosis Risk and Severity: A Mendelian Randomization Study. *medRxiv.* 2022 Jan 1.

Storm CS, Kia DA, **Almramhi M**, Bandres-Ciga S, Finan C, Hingorani AD,IPDGC, Wood NW. "Finding drug targeting mechanisms with genetic evidence for Parkinson's disease". Nat Commun 12, 7342 (2021). https://doi.org/10.1038/s41467-021-26280-1

Storm CS, Kia DA, **Almramhi M**, Athauda, IPDGC, Burgess S, Foltynie T, Wood NW. "Predicting the efficacy of exenatide in Parkinson's disease using genetics – a Mendelian randomization study". *medRxiv*. 2020.

Coneys R, Storm CS, Kia DA, **Almramhi M**, IPDGC, Wood NW. "Mendelian randomization finds no association between urate and Parkinson's disease progression". *Movement Disorders*. 2021.

Poster and achievements arising from this work

- Almramhi M ,Wood NW. “Investigating the Causal Relation Among plasma lipid Levels and Multiple sclerosis : A Bidirectional Two-sample Mendelian Randomisation Study”. 2019. This was a poster presentation at the European Neuroscience Conference by Doctoral Students Conference.

  Spotting and fixing issues in TwoSampleMR R package

- In version 0.5.4 of TwoSampleMR R package, when converting to MRInput format and supplying an LD matrix, it is possible that multi-allelic variants will be represented differently on in the GWAS and the LD reference panel. These ambiguous alignments were not being removed. I was able to spot and fix this issue. https://mrcieu.github.io/TwoSampleMR/news/index.html

- In version 0.5.2 of TwoSampleMR R package, when converting to MRInput format and supplying an LD matrix, the LD matrix SNP order was not matching the summary data order. I was able to spot and fix this issue. https://mrcieu.github.io/TwoSampleMR/news/index.html

Table of Contents

Abstract ... 3
Impact Statement ... 4
Acknowledgements ... 6
Publications ... 8
Poster and achievements arising from this work ... 9
List of Figures ... 14
List of Tables ... 16
Abbreviations ... 17
Chapter 1 Introduction ... 20
1.1. Introduction to multiple sclerosis ... 20
1.2. Immune pathogenesis of multiple sclerosis ... 22
1.3. Pathophysiology of multiple sclerosis ... 25
1.4. Epidemiology ... 28
1.5. Aetiology of multiple sclerosis ... 30
1.6. Comorbidity in multiple sclerosis ... 31
1.7. Diagnosis of multiple sclerosis ... 32
1.8. Disability measures in multiple sclerosis ... 34
1.9. Treatment for multiple sclerosis ... 36
1.10. Genome-wide association studies in multiple sclerosis ... 39
1.11. Mendelian randomisation approach ... 40
1.11.1. Mendelian randomisation for causal inference ... 41
1.11.2. Mendelian randomisation for drug repurposing ... 41
1.12. Mendelian randomisation in multiple sclerosis ... 44
1.13. Thesis motivation and aim ... 45
1.13.1. Specific objectives of this thesis ... 46
Chapter 2 Method ... 47
2.1. Overview ... 47
2.2. Dataset identification and summary statistics extraction ... 50
2.2.1. GWAS datasets descriptions overview ... 50
2.2.2. Exposure datasets ... 52
2.2.2.1. Lipid fractions data ... 52
2.2.2.2. Expression quantitative trait loci data ... 52

2.2.2.3. Neale Lab UK Biobank GWAS 53
2.2.2.4. GWAS dataset for stroke 54
2.2.3. MS outcome data 55
2.3. MR analysis 56
2.3.1. Prioritising and validating the assessment of genetic variants 56
2.3.1.1. Identify the genetic instruments 56
2.3.1.2. Independence assessment 56
2.3.1.3. Strength of the genetic instrument assessment 57
2.3.1.4. Harmonising genetic instruments across the dataset 58
2.3.1.5. Directionality assessment 59
2.3.2. Estimating the causal effect 60
2.3.2.1. Overview 60
2.3.2.2. Inverse-variance weighted 62
2.3.2.3. Sensitivity analyses 63
2.3.2.3.1. MR–Egger method 63
2.3.2.3.2. Weighted median 64
2.3.2.3.3. Pleiotropic assessment 65
2.3.2.3.3.1. Cochran's Q and I^2 tests 65
2.3.2.3.3.2. MR-Egger intercept 66
2.3.2.3.3.3. IVW MR-Radial 66
2.3.2.4. Further Analysis 67
2.3.2.4.1. Multivariable MR analysis 67
2.3.2.4.2. Bidirectional MR 70
2.3.2.5. Causal effect estimation using correlated variants 70
2.3.2.6. Multiple test correction 71
2.3.2.7. Reporting the findings 71
2.3.2.8. Specialist software for MR statistical analyses 72
Chapter 3 Exploring the Role of Lipids and Statins on Multiple Sclerosis Risk and Severity 73
3.1. Introduction 74
3.1.1. Limitation of genetically mimicking the effects of statins via HMGCR 75
3.1.2. Statin activity via the cholesterol-dependent pathway 77
3.1.3. Statin activity via the cholesterol-independent pathway (pleiotropic effects via Rho GTPases) 78
3.1.4. Effects of statins mediated by Rho GTPases 82
3.2. Aims and hypothesis 83
3.3. Method and results 85
3.3.1. Genetically predicted HDL-C associated with increased MS risk but not severity 86
3.3.2. Genetically predicted MS risk not associated with lipid levels (reverse causation analysis) ... 93
3.3.3. MR estimates for mimicking the effects of statins using expression data 96

3.3.3.1. Genetically mimicked effect of statins on MS risk is independent of cholesterol....................98
3.3.3.2. Genetically mimicked effect of statins had no causal association with MS severity 102
3.4. Discussion 105
3.4.1. High plasma HDL-C is a risk factor for MS 105
3.4.2. The lower cholesterol levels induced by statins have no effect on MS risk 107
3.4.3. The effects induced by statins via the cholesterol-independent pathway (*RAC2*) may reduce MS risk 108
3.4.4. Plasma lipids have no causal link to MS severity....................110
3.4.5. Genetically mimicked effects of statins have no causal links to MS severity 111
3.4.6. limitations 112
3.4.7. Conclusion 113
Chapter 4 Applying Mendelian randomization to identify potential druggable targets for multiple sclerosis 114
4.1. Introduction 115
4.2. Aims 117
4.3. Method 118
4.4. Results 120
4.4.1. MS's disease risk – discovery and replication....................121
4.4.2. MS severity....................127
4.4.3. Discussion 128
4.4.4. Limitation 134
4.4.5. Conclusion 135
Chapter 5 The role of body fat in multiple sclerosis susceptibility and severity 138
5.1. Introduction 139
5.2. The role of body fat in multiple sclerosis....................141
5.3. Aims 142
5.4. Method 143
....................146
5.4.1. Influence of genetically raised anthropometric measures on the risk of MS 147
5.4.2. Influence of genetically raised MS risk on anthropometric measures 150
5.4.3. Influence of genetically raised anthropometric measures on MS severity 152
5.5. Discussion 154
5.5.1. limitations 155
5.5.2. Conclusion 158
Chapter 6 Stroke as a risk factor for MS....................159
6.1. Introduction 160
6.2. Aims 162
6.3. Method 163
6.4. Results 164

6.4.1. No causal link between genetically predicted MS susceptibility and the risk of stroke 165
6.4.2. Causal link between genetically predicted IS and MS risk 169
6.4.3. No causal link between genetically predicted stroke and MS severity 174
6.5. Discussion 175
6.5.1. Limitation 178
6.5.2. Conclusion 180
Chapter 7 Discussion 181
7.1. Highlighting the key limitations of genetic and non-genetic studies on informed causal inference and drug repurposing 181
7.2. Mendelian randomisation 182
7.3. Pipeline for Mendelian randomisation analysis 182
7.4. Mendelian randomisation for drug target repurposing 183
7.5. Mendelian randomisation for causal analysis of risk factors 184
7.6. Future aims for MS severity and clinical courses 185
7.7. Future directions of Mendelian randomisation 185
Chapter 8 Appendix 187
Bibliography 197

List of Figures

Figure 1.1: Multiple sclerosis disease course……………………………………………………23

Figure 1.2: Possible mechanism of cross-reactivity between self-antigens and pathogen agents….24

Figure 1.3: Immune pathogenesis of multiple sclerosis……………………………………….26

Figure 1.4: The mechanisms of MS pathophysiology…………………………………………29

Figure 1.5: The Kurtzke expanded disability status scale………………………………………37

Figure 1.6: Schematic comparison between a randomised control trial (RCT) and Mendelian randomised (MR) study……………………………………………………………………..44

Figure 2.1: Schematic diagram illustrating the three key assumptions for MR…………………….49

Figure 2.2: Schematic diagram outlining the standard analysis workflow for an MR study in general and for this thesis specifically…………………………………………………………….50

Figure 2.3: A flowchart illustrating how the MR results have been interpreted in the present thesis…………………………………………………………………………………………62

Figure 2.4: A detailed flow chart showing the steps involved in genetic instrument selection and the interpretation of the results for MVMR analysis……………………………………………70

Figure 3.1: The cholesterol biosynthesis pathway. 3-Hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR) inhibition by statins reduces the synthesis of cholesterol and isoprenoids. …79

Figure 3.2: The regulation of Rho proteins……………………………………………………81

Figure 3.3: A flow diagram summarising this chapter's method and results………………………86

Figure 3.4: Scatter plots for MR analyses showing the causal estimates of the lipid fractions on MS risk……………………………………………………………………………………………90

Figure 3.5: Scatter plots for MR analyses showing the causal estimates of the lipid fractions on MS severity……………………………………………………………………………………….93

Figure 3.6: Scatter plots for MR analyses showing the causal estimates of the MS risk on lipid fractions……………………………………………………………………………………….96

Figure 3.7: Scatter plots for MR analyses showing the causal estimates of RAC2 on MS risk….103

Figure 4.1: Workflow and summarised results of the current study. ………………………………121

Figure 4.2: The forest plot visualises the discovery phase results for the 14 replicated genes……..123

Figure 4.3: Scatter plots of causal estimates of druggable genes that did not pass the sensitivity tests for MS risk. ………………………………………………………………………………125-127

Figure 5.1: Genetic correlations between the 21 anthropometric measures. ………………………146

Figure 5.2: Direct acyclic graph Adapted collider bias in case-only studies………………………157

Figure 6.1: A scatter plot relating the effect sizes of the SNP–MS risk associations and the SNP–IS risk associations……………………………………………………………………………..167

Figure 6.2: A scatter plot relating the effect sizes of the SNP–IS associations and the SNP–MS risk associations……………………………………………………………………………………168

Figure 6.3: A scatter plot displaying the effect sizes of the SNP–IS association and the SNP–MS risk associations……………………………………………………………………………………172

Figure 6.4: A scatter plots displaying the effect sizes of the SNP–AS and SNP–IS-subtype associations and the SNP–MS risk association. ……………………………………………………173

List of Tables

Table 1.1: Worldwide prevalence of multiple sclerosis per 100,000 people in 2013 and 2020 29
Table 1.2: 2017 McDonald criteria for the diagnosis of multiple sclerosis 33
Table 1.3: Differential diagnosis of multiple sclerosis 34
Table 1.4: Summary of approved disease-modifying therapies for multiple sclerosis......................................37
Table 2.1: Summary of GWAS used in the thesis. 51
Table 2.2: The methods used to estimate and evaluate the causal effect estimate 60
Table 3.1: Members of the Rho GTPase family 81
Table 3.2: Sample characteristics of the lipid traits 86
Table 3.3: MR analysis of the effect of lipid level on MS risk 88
Table 3.4: Replication analysis results for the effect of HDL-C on MS risk..90
Table 3.5: MR analysis of the effect of lipid level on MS severity 91
Table 3.6: MR analysis of the effect of MS risk on lipid levels 94
Table 3.7: The Rho GTPase family and mevalonate pathway (cholesterol biosynthesis pathway) gene list (*part 1*)..97
Table 3.8: MR estimates for the genetically mimicked effects of statins on MS risk (*part 1*) 100
Table 3.8: MR estimates for the genetically mimicked effects of statins on MS risk (*part 2*) 101
Table 3.9: MR estimates for the genetically mimicked effects of statins on MS severity (*part 1*) 103
Table 4.1: MR results on effects of druggable genes on MS risk after excluding outliers.............................127
Table 4.2: Summary of the existing drugs that can target the prioritised genes for MS risk...........................136
Table 5.1. Sample characteristics for the traits of interest 145
Table 5.2: MR estimates for the effect of anthropometric measures on MS risk (significance level is set at IVW FDR ≤ 0.05)...149
Table 5.3: MR estimates for the effect of MS risk on anthropometric measures (bidirectional analysis).......151
Table 5.4: MR estimates for the effect of anthropometric measures on MS severity (significance level is set at IVW FDR ≤ 0.05)..153
Table 6.1: GWAS datasets included in the MVMR analysis 164
Table 6.2: The estimated effects of MS on the risk of strokes based on MR 166
Table 6.3: The estimated effects of strokes on the risk of MS based on MR 171
Table 6.4: The estimated effects of stroke on MS severity based on MR 174

Abbreviations

Single nucleotide polymorphisms	SNP
Global Lipids Genetics Consortium	GLGC
Genotype-Tissue Expression	GTEx
Expression quantitative trait loci	eQTL
Megabase	Mb
Multiple sclerosis	MS
Central nervous system	CNS
Relapsing remitting	RRMS
Primary progressive	PPMS
Secondary progressive	SPMS
Myelin basic protein	MBP
T-cell receptor	TCR
Major histocompatibility complex	MHC
Blood brain-barrier	BBB
Antigen-presenting cells	APCs
Vascular cell adhesion molecule-1	VCAM-1
Intracellular cell adhesion molecule-1	ICAM-1
Tumour necrosis factor	TNF
Interferons	IFN
T helper	Th
Chemokine receptor 6	CCR6

C-C Motif Chemokine Ligand 20	CCL20
Natural killer cells	NK
Experimental autoimmune encephalomyelitis	EAE
Genome wide association studies	GWAS
International Multiple Sclerosis Consortium	IMSGC
Randomised controlled trials	RCTs
Mendelian randomization (MR)	MR
Low-density lipoprotein cholesterol	LDL-C
High-density lipoprotein cholesterol	HDL-C
Triglycerides	TG
Coronary heart disease	CHD
Expanded disability status scale	EDSS
Disease modifying therapies	DMTs
Ischemic stroke	IS
Large-artery atherosclerotic stroke	LAS
Cardioembolic stroke	CES
Small-vessel disease	SVS
Systolic blood pressure	SBP
Diastolic blood pressure	DBP
Confidence interval	CI
Inverse-variance weighted	IVW
Multivariable Mendelian randomization	MVMR
Body mass index	BMI

Chapter 1 Introduction

1.1. Introduction to multiple sclerosis

Multiple sclerosis (MS), the most common neurological disability, is an autoimmune-mediated, demyelinating disorder that affects the central nervous system (CNS) (Thompson A, 2018, Brownlee et al., 2017). MS is characterised by inflammation causing episodic attacks, multifocal demyelination and axonal loss, leading to changes in sensation, mobility, balance, sphincter function, vision, and cognition(Brownlee et al., 2017) . In 1996, the US National Multiple Sclerosis Society (NMSS) Advisory Committee on Clinical Trials in MS defined the clinical subtypes of MS into four different subtypes: relapsing-remitting (RRMS), secondary progressive (SPMS), primary progressive (PPMS), and progressive relapsing (PRMS) (Lublin et al., 2014) (**Figure 1.1**). *These demyelinated plaques consist of a defined* (Lublin et al., 2014).

- RRMS is the most common form, found in about 85% of MS patients, and is characterised by periods of relapse (flare-up of symptoms) or episodes of neurological dysfunction over days or weeks, followed by remission with full or partial periods of recovery over months or years (Mansilla et al., 2021, Lublin et al., 2014). During RRMS, inflammatory attacks on myelin and nerve fibres occur (Ghasemi et al., 2017). During these inflammatory episodes, activated immune cells cause various neurological symptoms, such as visual impairments, tingling and numbness, episodic bouts of fatigue, intestinal and urinary system disorders, spasticity and learning, memory impairment and weakness (Ghasemi et al., 2017).

- Nearly 15%−30% of patients with RRMS eventually progress into SPMS, in which there is a steadily progressive worsening of neurologic function (accumulation of disability) over time, with or without relapses (Thompson A, 2018, Lublin et al., 2014). It is difficult to determine exactly when the transition from RRMS to SPMS starts, but the research data suggest an average time to the progressive phase of about 19 years after the onset of RRMS (Klineova and Lublin, 2018). However, an earlier progression to SPMS is associated with a higher age at RR onset, male gender (albeit not consistently in all studies), spinal cord symptoms and incomplete relapse recovery (Klineova and Lublin, 2018). RRMS patients experience symptoms of increased weakness, intestinal and urinary system disorders, fatigue, stiffness, mental disorders , psychological impairment and sensory symptoms (Ghasemi et al., 2017).

- PPMS affecting approximately 10 % of all MS patients (Wolinsky et al., 2007) . Patients with PPMS experience steadily worsening neurologic function from onset without any distinct relapses or remissions (Wolinsky et al., 2007, Lublin et al., 2014, Klineova and Lublin, 2018). PPMS patients often experience symptoms that indicate a spinal cord progressive disease, including problems with walking, weakness, stiffness, and trouble with balance (Ghasemi et al., 2017).

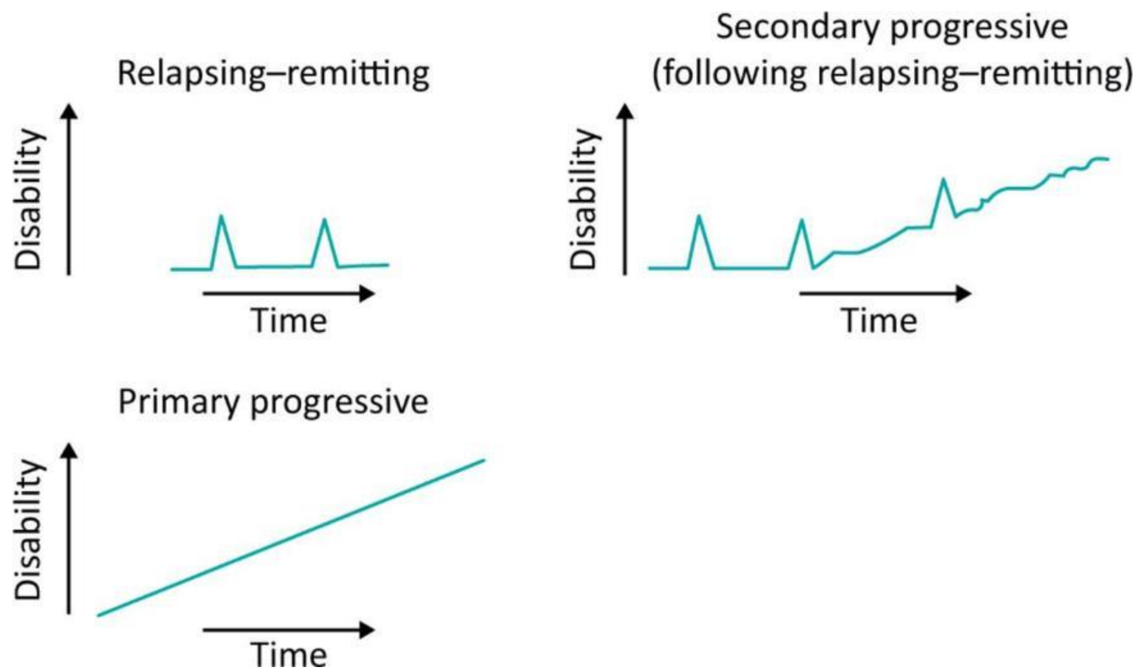


Figure 1. 1: Multiple sclerosis disease course adapted from Ford, 2020.

1.2. Immune pathogenesis of multiple sclerosis

Loss of immunological tolerance to self-antigens is the hallmark of autoimmune diseases, including MS. It has been shown that the immune system identifies CNS myelin as foreign and is subsequently activated to attack it (Schaeffer et al., 2015a). Two mechanisms have been proposed to explain the initiation of MS: molecular mimicry and bystander activation.

The molecular mimicry mechanism occurs when peptides of infectious agents share sequences or have structural similarities with self-antigens (Schaeffer et al., 2015a). This similarity may trigger activation of T and/or B cells, leading to a crossing of the blood-brain barrier (BBB) and tissue damage upon recognition of antigens in the brain (Schaeffer et al., 2015a). Many molecular mimics (bacteria and viruses) have been identified in MS studies (Libbey et al., 2007). For example, myelin basic protein (MBP), the major constituent of the myelin sheath of oligodendrocytes, is one autoantigen that shares

similar T-cell receptor (TCR) motifs and major histocompatibility complex (MHC)-II binding with viruses, such as herpes simplex virus (HSV) (Schaeffer et al., 2015a) (**Figure 1.2**).

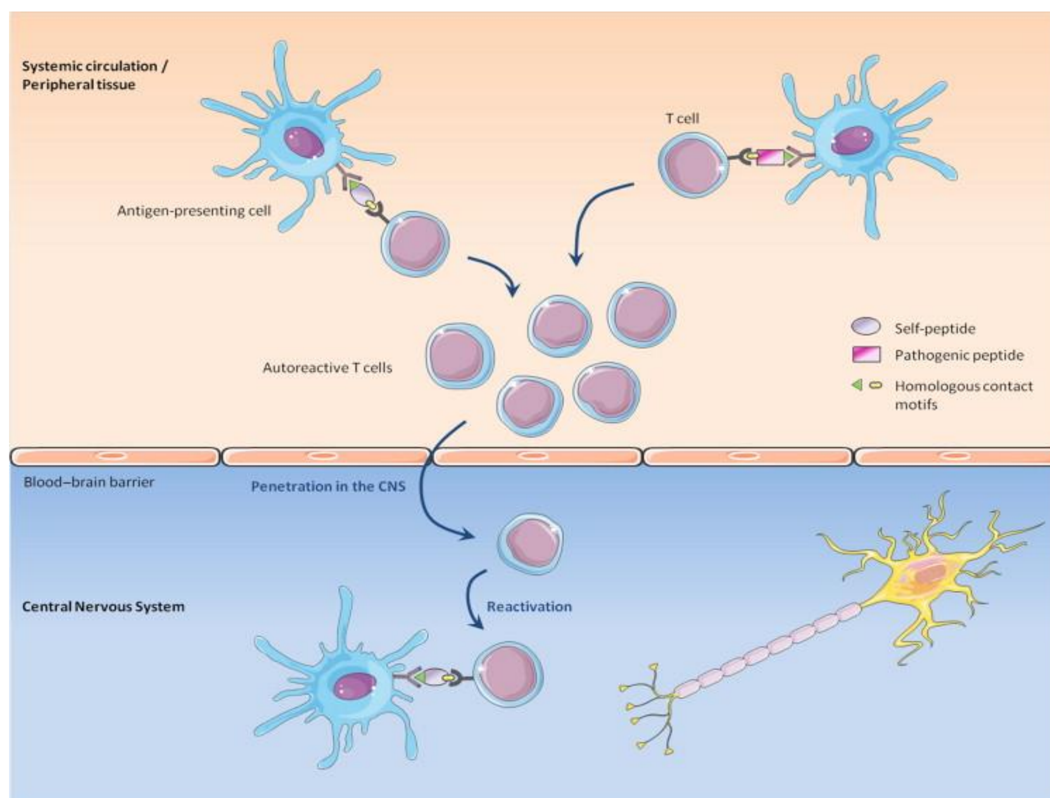


Figure 1. 2: Possible mechanism of cross-reactivity between self-antigens (e.g. myelin basic protein) and pathogen agents (e.g. Epstein–Barr virus, herpes simplex virus). This figure was reproduced from Schaeffer et al., 2015.

The second mechanism is bystander activation of T-cells, which occurs non-specifically during infections (in a TCR-independent manner) (Schaeffer et al., 2015a) (**Figure 1.3**). Classical T-cell activation occurs mainly by engagement of the TCR, which triggers several signalling cascades that result in cytokine production, differentiation, proliferation and/or apoptosis (Pacheco et al., 2019). In bystander activation of T-cells (TCR-independent manner), activation is mediated through indirect signals that favour an inflammatory milieu, such as inflammatory cytokines, superantigens and toll-like receptor activation (Schaeffer et al., 2015a, Pacheco et al., 2019). Bystander activation of autoreactive T-cells might also occur via APCs (Schaeffer et al., 2015a). In this process, virus

infections trigger the activation of APCs, such as dendritic cells, which then activate preprimed autoreactive T-cells to initiate the autoimmune response (Schaeffer et al., 2015a, Fujinami et al., 2006).

In addition, bystander activation of autoreactive T cells can also be initiated by virus-specific T-cell (depending on specific TCR recognition) (Schaeffer et al., 2015a, Fujinami et al., 2006, Pacheco et al., 2019). For example, activated virus-specific T-cells migrate to the infected CNS, where the virus-infected cells present viral peptides in the context of MHC to T-cells (Schaeffer et al., 2015a, Fujinami et al., 2006, Pacheco et al., 2019). These immune cells recognise the infected cells and release cytokines that kill the infected cells (Schaeffer et al., 2015a, Fujinami et al., 2006, Pacheco et al., 2019). This results in the secretion of inflammatory cytokines, such as tumour necrosis factor (TNF), lymphotoxin and nitric oxide, by dying cells, leading to the destruction of the uninfected neighbouring cells and the release of autoantigens (Schaeffer et al., 2015a, Fujinami et al., 2006, Pacheco et al., 2019). Under infectious circumstances, the presentation of these autoantigens may activate autoreactive T-cells (Schaeffer et al., 2015a, Fujinami et al., 2006, Pacheco et al., 2019).

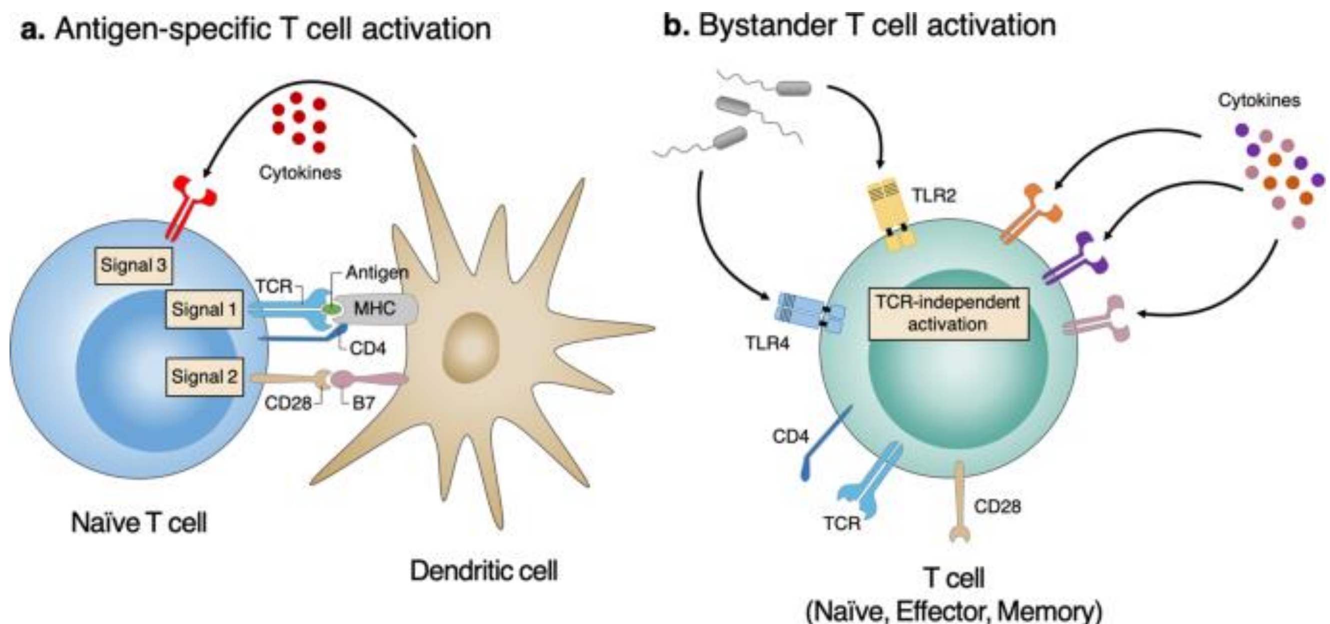


Figure 1.3: Immune pathogenesis of multiple sclerosis**. (a)** Antigen-specific T-cell activation (classic T-cell activation) requires three distinct signals. Signal 1 is antigen-specific signalling generated due to the engagement of the T-cell receptor with pathogenic peptides presented by major histocompatibility complex molecules. Signal 2 is costimulatory signalling generated due to the interaction of CD28 with one of the B7 molecules (CD80 and CD86). Signal 3 is polarising signalling generated due to the various cytokine milieus produced by dendritic cells. (**b)** On the other hand, bystander T-cell activation is the concept of T-cell activation independent of antigen stimulation. Bystander-activated T-cells can respond rapidly to inflammatory mediators, such as cytokine and Toll-like receptors, signalling in a T-cell receptor-independent manner (Lee et al., 2020a). TLR2: Toll-like receptor 2, TLR4: Toll-like receptor 4. This figure was reproduced from Lee et al., 2020.

1.3. Pathophysiology of multiple sclerosis

The pathological hallmark of MS is the presence of demyelinated plaque or lesion within the CNS (grey and predominantly white matter) (Schaeffer et al., 2015a). These demyelinated plaques consist of a defined hypocellular area characterised by the loss of myelin and the formation of an astrocytic scar (Schaeffer et al., 2015a). Evidence has also established that transected axons, a consistent consequence of demyelination, are common in the plaques of MS, and axonal transection might be the pathologic correlate of the irreversible neurologic impairment in MS (Trapp et al., 1998).
Lesions are found in different areas within the CNS, including the optic nerves, periventricular white matter, brainstem, cerebellum and spinal cord white matter, and they often surround medium-sized blood vessels (Schaeffer et al., 2015a). A cascade of pathobiological events characterises the pathological process of MS, ranging from autoreactive T-cell activation and breakdown of BBB to demyelination and axonal degeneration (**Figure 1.4**).

In the periphery, autoreactive CD^{4+} T cells are activated by APCs, which present via the MHC class II receptor, an amino acid similar to myelin peptides synthesised in the CNS (Celarain and Tomas-Roig, 2020). This interaction activates the differentiation of the CD^{4+} T naïve cells into CD^{4+} T helper (Th)

cells (Celarain and Tomas-Roig, 2020). Upon activation, the Th1 subtypes release pro-inflammatory cytokines (IL-1, IFN- γ, TNF, IL-12, IL-17), which initiate different events, including recruiting other pro-inflammatory cells, such as cytotoxic CD^{8+} T, B-cells and monocytes, expressing ligands (i.e. integrins, chemokine receptor 6; CCR6) on the T-cell surface and adhesion molecules of BBB endothelial cells (i.e. C-C Motif Chemokine Ligand 20; CCL20) (Cervantes-Gracia and Husi, 2018, Celarain and Tomas-Roig, 2020). These proinflammatory cells then migrate across the disrupted BBB through the endothelium and the endothelial basal lamina into the CNS parenchyma (Schaeffer et al., 2015a, Celarain and Tomas-Roig, 2020).

The activation and recruitment of leukocytes lead BBB endothelial cells to reorganise their membrane in multiple cup-shaped microdomains enriched with cellular adhesion molecules, such as vascular cell adhesion molecule-1 (VCAM-1) and intracellular cell adhesion molecule-1 (ICAM-1) (Schaeffer et al., 2015a). These molecules surround the migrating cells to enable their passage across the endothelium cells (Schaeffer et al., 2015a). Under healthy conditions, the brain is protected by intact BBB, which tightly regulates the passage of substances from the blood circulation to the brain and orchestrates immune surveillance of the CNS (Celarain and Tomas-Roig, 2020). In MS, MRI clinical observations of acute and chronic active lesions reveal a disruption in the BBB, which enables the infiltration of lymphocytes and leukocytes into the CNS (Schaeffer et al., 2015a) . Indeed, the release of proinflammatory cytokines, such as IFN-γ and TNF-α, contributes to BBB disruption by activating cerebral endothelial cells and modulating the BBB phenotype through the induction of several inflammatory genes (Celarain and Tomas-Roig, 2020, Schaeffer et al., 2015a). These genes affect BBB integrity through several mechanisms, such as the inhibition of junctional protein expression, resulting in massive lymphocyte trafficking into the brain (Celarain and Tomas-Roig, 2020, Schaeffer et al., 2015a). Activated T-cells that migrate across the BBB also affect BBB integrity by expressing matrix metalloproteinases, which contribute to BBB breakdown (Celarain and Tomas-Roig, 2020, Schaeffer et al., 2015a). Once CD^{4+} T cells are in the CNS, they reactivate upon interaction with the resident APCs, leading to the release of a variety of proinflammatory cytokines and chemokines that result in

astrogliosis and microgliosis (Celarain and Tomas-Roig, 2020, Schaeffer et al., 2015a). The infiltrated cytotoxic CD8+ T cells attack oligodendrocytes, causing their destruction and neuronal death (Celarain and Tomas-Roig, 2020). The B-cells contribute to myelin destruction through several mechanisms, including complement-mediated opsonisation, which facilitates phagocytosis by macrophages, complement-mediated cytolysis or stimulation of antibody-dependent cell-mediated cytotoxicity by binding to natural killer (NK) cells (Schaeffer et al., 2015a).

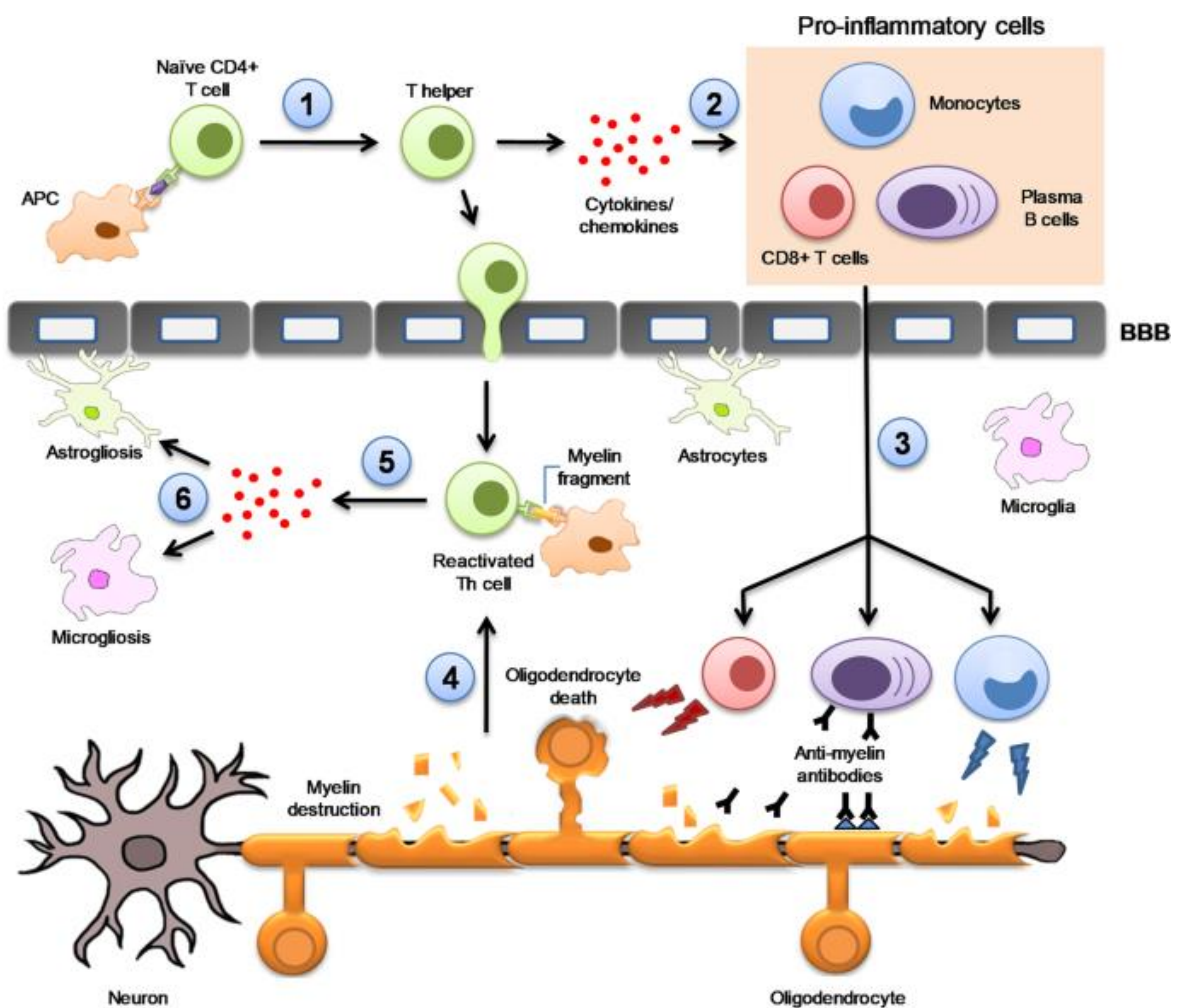


Figure 1.4: The mechanisms of MS pathophysiology reproduced from (Celarain and Tomas-Roig, 2020). In the periphery, autoreactive CD^{4+} T cells are activated by antigen-presenting cells (APCs) that present, in conjunction with class II MHC molecules, antigens similar to myelin. (1) Following this interaction, CD4+ T naïve cells differentiate into CD^{4+} T helper cells (Celarain and Tomas-Roig, 2020). (2) These cells subsequently release proinflammatory cytokines (such as interferon-gamma; IFN-γ), which recruit other pro-inflammatory cells, such as cytotoxic CD^{8+} T, B-cells and macrophages (Celarain and Tomas-Roig, 2020). (3) These proinflammatory cells migrate to the blood-brain barrier and pass into the CNS (Celarain and Tomas-Roig, 2020). Inside the brain, plasma B-cells generate autoantibodies against CNS self-antigens, contributing to myelin sheath damage (Celarain and Tomas-Roig, 2020). This process is exacerbated when infiltrated cytotoxic CD^{8+} T cells attack oligodendrocytes, causing their destruction and neuronal death. Monocytes, on the other hand, contribute to demyelination through myelin phagocytosis (Celarain and Tomas-Roig, 2020). (4) In parallel, infiltrated CD^{4+} T cells are reactivated upon interaction with myelin fragments presented by resident APCs, leading to (5) proinflammatory cytokine and chemokine release and (6) astrogliosis and microgliosis (Celarain and Tomas-Roig, 2020).

1.4. Epidemiology

MS usually starts between the ages of 20 and 40 years, but approximately 1–4% of cases can occur in childhood, and around 2–10% occur after 50 years of age (Ghasemi et al., 2017, Schaeffer et al., 2015a). There is wide geographical variance in the incidence and prevalence of MS depending on gender, age, geographic distribution and ethnic origin (Simpson et al., 2011, Schaeffer et al., 2015a). Women are more susceptible to MS compared with men; the prevalence is three times higher in females than in males (Schaeffer et al., 2015a). The reasons for this are not fully understood, but one explanation has been suggested: the possible involvement of sex hormones in susceptibility to MS (Harbo et al., 2013, Schaeffer et al., 2015a). It has been shown that high levels of progesterone, oestradiol and oestriol during pregnancy ameliorate the disease course of MS (Harbo et al., 2013). Further, in the animal model of MS experimental autoimmune encephalomyelitis (EAE), a milder disease course was observed under the influence of estrogen (Harbo et al., 2013).

Approximately 2.8 million people worldwide are affected by MS (Walton et al., 2020). Regionally, Europe has the greatest MS prevalence, followed by the Americas, whereas it is lowest in Southeast Asia and Africa (Walton et al., 2020). The higher MS incidence in some regions, such as European countries, is possibly due to living in countries at higher latitudes, resulting in reduced exposure to sunlight and thereby low vitamin D levels (Tao et al., 2016, Wood, 2017).

Table 1. Table 1.1 shows the number of people living with MS worldwide in 2013 and 2020, with the indication that the MS prevalence in 2020 is 30% higher than that in 2013 (Walton et al., 2020). The trend in MS prevalence could be related to many factors, including the greater awareness of the disease, increased accessibility to high-quality health care, earlier diagnosis, and the long-term survival of patients with MS (Walton et al., 2020).

Table 1.1: Worldwide prevalence of multiple sclerosis per 100,000 people in 2013 and 2020

	Number of countries included	**2013 prevalence per 100,000 population [95% CI]**	**2020 prevalence per 100,000 population [95% CI]**	**Increase; absolute (%)**
Global	81	29.26 [29.21, 29.30]	43.95 [43.90, 44.01]	14.69 (50%)
African	6	5.52 [5.41, 5.62]	8.76 [8.64, 8.89]	3.24 (59%)
Americas	15	62.89 [62.72, 63.05]	117.49 [117.27, 117.71]	54.6 (87%)
E. Mediterranean	14	23.91 [23.77, 24.04]	33.00 [32.85, 33.15]	9.09 (38%)
European	35	108.25 [108.01, 108.49]	142.81 [142.53, 143.08]	34.56 (32%)
South East Asia	4	5.44 [5.41, 5.48]	8.62 [8.58, 8.66]	3.18 (58%)
Western Pacific	7	3.64 [3.61, 3.67]	4.79 [4.75, 4.82]	1.15 (32%)

This table is adapted from Walton et al., 2020. Abbreviations: CI, confidence intervals; E. Mediterranean, Eastern Mediterranean.

1.5. Aetiology of multiple sclerosis

Although the aetiology of MS remains unclear, the present evidence suggests that the cause of MS is multifactorial, including genetic and environmental determinants (Jacobs et al., 2021, Ghasemi et al., 2017). A large fraction of MS risk is explained by lifestyle and environmental factors, while genetic predisposition to MS only explains a small fraction (Olsson et al., 2017). Environmental factors include exposure to smoking, low vitamin D levels due to insufficient sun exposure and/or dietary intake, obesity, viral and bacterial agents, such as Epstein-Barr virus, organic solvents and shift work (Olsson et al., 2017). Interestingly, these nongenetic factors can influence pathogenetic mechanisms, and some of them can be modified (Olsson et al., 2017). For example, a recent longitudinal analysis comprising more than 10 million young adults on active duty in the US military, 955 of whom were diagnosed with MS during their period of service, tested the hypothesis that MS is caused by Epstein–Barr virus (Bjornevik et al., 2022). The results revealed that the risk of MS prevalence increased 32-fold after infection with Epstein–Barr virus but was not increased after infection with other viruses, including the similarly transmitted cytomegalovirus (Bjornevik et al., 2022). This finding suggests that Epstein–Barr virus could be a potential modifier of MS.

Compelling evidence supports the important role of genetic determinants in MS aetiology. Thus far, more than 100 candidate genes within and out of MHC regions have been found to be associated with MS risk. Different strategies have been used to identify these candidate genes, including population-based association studies, family-based linkage methods and systematic genome screens (Olsson et al., 2017).

Epidemiological data also suggest an interaction between genetic predispositions and environmental factors, whereby the effect of certain loci may depend on exposure to environmental risk factors (Jacobs et al., 2021, Olsson et al., 2017). For example, evidence from a population-based case-control study suggested that smoking's influence on MS risk can be modified by the human leukocyte antigen (HLA) genotype (Hedström et al., 2011). This study found that smokers carrying HLA-DRB1*15 and

lacking HLA-A*02 had a 14-fold (odds ratio 13.5, 95% CI 8.1–22.6) increased risk compared to non-smokers without these genetic risk factors (odds ratio 4.9, 95% CI 3.6–6.6) (Hedström et al., 2011).

1.6. Comorbidity in multiple sclerosis

Comorbidity, defined as the total burden of illness other than the disease of interest (MS), in patients with MS has drawn much interest in recent years due to its high prevalence and breadth of adverse impacts (Marrie, 2017). Several well-designed population-based studies have sought to assess the prevalence of comorbidities among patients with MS. For example, in a recent systematic review, *Marrie et al.* reported the incidence and prevalence of the most common comorbidities in MS, including depression, anxiety, hypertension, hyperlipidaemia, and chronic lung disease (Marrie et al., 2015) . In addition, other studies have investigated the incidence or prevalence of many autoimmune diseases, cancers, ischaemic heart disease, and stroke among patients with MS (Marrie et al., 2015). Together, these studies shed light on how these comorbid diseases adversely affect MS by accelerating disability progression, increasing changes visible on MRI, increasing mortality, reducing quality of life, and delaying diagnosis (Marrie and Horwitz, 2010, Marrie, 2017, Magyari and Sorensen, 2020). They are also improved our knowledge awareness of the comorbidity prevention or treatment using lifestyle modification and readily available treatments (Marrie, 2017).

The reason for the co-occurrence of diseases is unknown, but it is likely true that common genes and/or similarities in immunologic features contribute to the susceptibility of an individual to more than one disease (Nielsen et al., 2006, Marrie and Horwitz, 2010). Common risk factors can also lead to an increased co-occurrence of diseases, although common environmental factors, such as obesity and smoking, are another possible explanation (Marrie and Horwitz, 2010). Furthermore, diseases might co-occur due to one of several aetiological mechanisms, including associated risk factors, direct causation, heterogeneity, and independence (Marrie and Horwitz, 2010).

1.7. Diagnosis of multiple sclerosis

There is no exact measure or laboratory marker for the diagnosis of MS. The diagnosis of MS is based on the integration of clinical, imaging, and laboratory findings, such as those from magnetic resonance imaging (MRI), lumbar punctures for cerebrospinal fluid (CSF) analysis, evoked potentials, and blood sample analysis (Thompson A, 2018). Clinical expertise is necessary to obtain evidence of dissemination in time and space and, importantly, to exclude other neurological conditions (Thompson A, 2018). MRI is the most sensitive test for assisting in excluding other conditions, permitting earlier diagnosis at an increased certainty with successive versions of the diagnostic criteria (Thompson A, 2018).

The McDonald criteria are employed to diagnose MS. These criteria were established by the International Panel on Diagnosis of Multiple Sclerosis in 2001 and revised several times in 2017 (**Table 1.2**). The goal of the criteria is to ensure an earlier and more accurate diagnosis of MS and minimise the uncertainty period for the patient and clinician (Ford, 2020). This provides appropriate management, including confirmation of the diagnosis for the patient and access to effective therapies (Ford, 2020).

Table 1.2: 2017 McDonald criteria for the diagnosis of multiple sclerosis

Number of attacks at clinical presentation	Number of lesions with objective clinical evidence	Additional data needed for diagnosis of multiple sclerosis
≥2	≥2	None[a]
≥2	1 (as well as clear-cut historical evidence of a previous attack involving a lesion in a distinct anatomical location)	None[a]
≥2	1	Dissemination in space demonstrated by an additional clinical attack implicating a different CNS site
		Or by MRI
1	≥2	Dissemination in time demonstrated by an additional clinical attack
		Or by MRI
		Or demonstration of CSF-specific oligoclonal bands
1	1	Dissemination in space demonstrated by an additional clinical attack implicating a different CNS site
		Or by MRI
		And dissemination in time demonstrated by an additional clinical attack
		Or by MRI
		Or demonstration of CSF-specific oligoclonal bands

[a] = No additional tests are needed to demonstrate dissemination in space and time. However, unless MRI is not possible, brain MRI must be obtained in all patients in whom a diagnosis of multiple sclerosis is being considered. In addition, spinal cord MRI or CSF examination should be considered in patients with insufficient clinical and MRI evidence supporting multiple sclerosis, with a presentation other than a typical clinically isolated syndrome or with atypical features. If imaging or other tests (e.g. CSF) are undertaken and are negative, caution needs to be taken before making a diagnosis of multiple sclerosis, and alternative diagnoses should be considered. Abbreviations: CNS, central nervous system; CSF, cerebrospinal fluid; MRI, magnetic resonance imaging. This table is adapted from (Ford, 2020).

Many other conditions have symptoms similar to those of MS. Therefore, clinicians need to be vigilant of atypical clinical findings or investigation results (Ford, 2020). One approach to solving such issues is the differential diagnosis of MS, which is the process of discovering the cause of symptoms by ruling

out other possible causes. The differential diagnosis of MS is wide and varies depending on the site of presentation, for instance, the optic nerve or spinal cord (**Table 13**) (Ford, 2020). Non-specific symptoms with white matter lesions on MRI can be a common cause of misdiagnosis of common disorders, such as migraine or small vessel vascular disease in the elderly (Ford, 2020).

Table 1.3: Differential diagnosis of multiple sclerosis

Autoimmune/inflammatory	CNS infections	Metabolic	Vascular conditions	Other
Neuromyelitis optica spectrum disorder (NMOSD)	CNS syphilis	Vitamin B_{12} deficiency	Small vessel disease	CNS lymphoma
Acute disseminated encephalomyelitis (ADEM)	Lyme disease	Copper deficiency	Stroke	Paraneoplastic
Myelin oligodendrocyte glycoprotein (MOG) antibody disease	Human T-lymphotropic virus (HTLV)	Mitochondrial disease	CADASIL	
Sjögren's syndrome	HIV	Leukodystrophies	Susac's syndrome	
CNS lupus			Anti-phospholipid antibody syndrome	
Sarcoidosis				
Behçet's				
CNS vasculitis				

This table is adapted from (Ford, 2020). Abbreviations: CADASIL, cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy; CNS, central nervous system.

1.8. Disability measures in multiple sclerosis

MS is a leading cause of disability. Therefore, in measuring disability, it is crucial to determine a patient's current and potential future severity to understand whether interventions affect the disease course and mitigate disease progression (Manouchehrinia et al., 2021).

The commonly used scales for assessing the level of disability in clinical practice include the expanded disability status scale (EDSS), MS severity score and age-related MS severity score (Manouchehrinia et al., 2021, Kurtzke, 1983). The EDSS is a measure that rates the impairment in MS patients on a scale ranging from 0 (no neurological impairment) to 10 (death as a result of MS) and comprises an assessment of eight functional systems conducted by a neurologist during a standard clinical examination (Kurtzke, 1983) (**Figure 1.5**). The EDSS is widely used as an outcome variable in MS clinical trials, possibly because it is easy to administer and often applied in clinical practice to monitor patients' progression over time (Manouchehrinia et al., 2017). However, evidence indicates that EDSS is not linear (Twork et al., 2010). This means, for example, that the difference between patients who walked with aid and wheelchair-bound patients was smaller than the difference between patients walking without help and patients being dependent on a walking aid in most health-related quality of life domains (Twork et al., 2010).

The is an algorithm that relates scores on the EDSS to the distribution of disability in patients with comparable disease durations (Roxburgh et al., 2005). The MS severity score is designed to predict disease severity. For example, if the patient accumulated disability at a faster-than-average rate compared to the patients with similar disease duration, then he/she experienced rapid disease progression (severe MS) (Kister and Kantarci, 2020). By contrast, the inverse is true if the patient has a lower-than-average disability relative to their peers with similar disease duration; in that case, the patient is classified as having mild MS (Kister and Kantarci, 2020). The MS severity score is widely used in different settings and has been shown to have a better statistical power to identify differences in disability between groups of patients than the other available measures of disease progression (Manouchehrinia et al., 2017).

However, the only drawback of this scale is its dependence on the date of disease onset, which is commonly assigned retrospectively and often missing, imprecise or unobtainable, resulting in a loss of data and subsequently loss of statistical power (Manouchehrinia et al., 2017).

The age-related MS severity score is another disability measure ranking EDSS scores based on the patient's age at the time of assessment (Manouchehrinia et al., 2017). It has been shown that age-related MS severity score is a more versatile tool and could minimise study biases and loss of statistical power caused by inaccurate or missing onset dates (Manouchehrinia et al., 2017). This is because this scale uses an individual's age instead of the onset date (Manouchehrinia et al., 2017, Manouchehrinia et al., 2021).

Figure 1.5: The Kurtzke expanded disability status scale. Reproduced from (Leddy and Dobson, 2020).

1.9. Treatment for multiple sclerosis

Currently, there is no cure or preventive measure for MS, and the approved treatments are disease-modifying therapies (DMTs) aimed at modifying the course of MS by slowing down disease progression, reducing relapses, decreasing long-term neurologic dysfunction, managing symptoms and, in some cases, modestly improving disability (Baecher-Allan et al., 2018, Hauser and Cree, 2020).

Table 1.4: Summary of approved disease-modifying therapies for multiple sclerosis

Drug (Date of Approval) *Commercial Name*	Target	Mechanism of Action	Efficacy	Adverse Events	Authorisation
IFNβ (1993) (IFNβ-1b *Betaseron, Extavia*) (IFNβ-1a *Avonex, Rebif, Plegridy*)	Binds to type I IFN receptor on human cells	Inhibits T cell division, matrix metallo-proteinase, BBB migration, and proinflammatory cytokines. Induces Tregs, suppressive transitional $CD19^{+}CD24^{++}$ $CD38^{++}$ B cells	Reduction in annualized relapse rate in RRMS. Delays conversion to clinically definite MS in CIS.	Post-injection flu-like symptoms, liver toxicity, and depression. Neutralizing Abs associated with reduced efficacy	Approved
Glatiramer Acetate (1997) ***Copaxone, Glatopa***	Random polymers of glutamic acid, lysine, alanine, and tyrosine to bind to MHC	Competes with peptide binding to MHC, increases IL- 10, IL-4, TGFβ, and CD8 Tregs, induces type II monocytes that cause a Th1 to Th2 shift	Reduction in annualized relapse rate in RRMS.	Post-injection idiosyncratic reactions, lymphadenopathy	Approved
Mitoxantrone (2000) ***Novantrone***	Interferes with DNA repair	Causes nucleotide crosslinking and DNA strand breaks. Inhibits lymphocyte and monocyte migration, B cell function, and secretion of TNFα, IL-2 and IFNγ	Reduction in annualized relapse rate and disease progression in clinically worsening RRMS and SPMS.	Bone marrow suppression, cardiomyopathy, leukemia	Not approved for use in MS in the UK,
Natalizumab (2004) ***Tysabri***	CD49d, the α4 subunit of VLA4 integrin (humanized mAb)	Blocks B and T cell migration into the CNS. Blocks VLA4 binding to VCAM-1 and fibronectin	Reduction in annualized relapse rate in RRMS.	Reactivation of the John Cunningham virus in the CNS may occur in some patients (PML), infusion reactions	Approved
Fingolimod (2010) ***Gilenya***	S1PR (Sphinosine-1-phosphate receptor)	Sequesters lymphocytes in lymph nodes by inhibiting lymphocyte egress. Inhibits the migration of dendritic cells to secondary lymphoid organs	Reduction in annualized relapse rate in RRMS.	Bradycardia, infection, macular edema, lymphopenia rare PML cases	Approved
Teriflunomide (2012) ***Aubagio***	Inhibits dihydroorotate dehydrogenase	Inhibits pyrimidine synthesis. Inhibits secretion of proinflammatory cytokines and T cell activation	Reduction in annualized relapse rate in RRMS.	Hair thinning, liver toxicity, teratogenesis	Approved
Dimethyl fumarate (2013) ***Tecfidera***	Nrf2 pathway	Activates the Nrf2 transcriptional pathway	Reduction in annualized relapse rate in RRMS.	GI side effects, flushing, lymphopenia, and rare PML	Approved
Alemtuzumab (2014) ***Lemtrada***	CD52 on T and B cells (humanized mAb)	Depletes B and T cells via ADCC and complement	Reduction in annualized relapse rate in RRMS.	Autoimmune diseases including thyroid, immune thrombocytopenia purpura, and glomerulonephritis	Approved
Daclizumab (2016) ***Zinbryta***	CD25, anti-IL2R (humanized mAb)	Prevents IL-2 signaling through the high affinity IL-2R. Augments CD56+ NK cell activity	Reduction in annualized relapse rate in RR.	Liver toxicity, skin reactions	* Withdrawn from the market in March 2018.
Ocrelizumab (2017) ***Ocrevus***	CD20+ B cells (humanized mAb)	Depletes CD20+ B cells. Reduces pathogenic B cell antigen presentation	Decreased annualized relapse rate in RRMS. Reduced disease progression in PPMS.	Infusion reactions, risk of non-melanoma skin cancer, infections, hypogammaglobulinemia.	Approved
Cladribine (2017)	Adenosine deaminase	Depletes immune cells by inducing lymphocyte apoptosis. Sustained reduction in CD4 and CD8 T cells and transient reduction in B cells	Reduction in annualized relapse rate in RRMS.	Infection, lymphopenia	Approved
Siponimod (2019) ***Mayzent***	sphingosine-1-phosphate (S1P) receptor modulator	Siponimod binds with high affinity to both S1P receptors 1 and 5 to block the ability of lymphocytes to release from the lymph nodes, decreasing the number of lymphocytes found in the peripheral blood.	For the treatment of relapsing forms of MS, to include clinically isolated syndrome, RRMS, and active SPMS in adults.	[a] Headache, back pain, bradycardia, dizziness, fatigue, influenza, urinary tract infection, lymphopenia, nausea and alanine amino transferase increases.	Approved
Ponesimod (2021) ***Ponvory***	sphingosine 1-phosphate receptor 1	Ponesimod modulates this response by stimulating and internalizing S1P1R on lymphocytes, effectively blinding them to concentration gradients of S1P, reducing the number of lymphocytes in blood	To treat adults with relapsing forms of MS, including CIS, RRMS, and active SPMS	[b] Anxiety, dizziness, dyspnoea, increased alanine aminotransferase, influenza, insomnia and peripheral oedema	Approved
Ofatumumab (2020) ***Arzerra, Kesimpta***	CD20	Ofatumumab binds to CD20, this binding allows ofatumumab to persist on the B lymphocyte cell surface for an extended period and recruit immunological molecules or FcR-expressing innate effectors, such as macrophages, that mediate immune effector functions with strong cytotoxic effects.	For the treatment of adult patients with relapsing forms of MS, including active SPMS, CIS, and RRMS	[c] Rash, erythema, upper respiratory tract infection, throat irritation, fatigue, headache, back pain and flushing.	Approved

This table is reproduced from (Baecher-Allan et al., 2018). Drug information for Siponimod, Ponesimod and Ofatumumab obtained from drug bank (https://go.drugbank.com/). Authorizing information collected from gov.uk (https://www.gov.uk/) Abbreviations: RRMS, relapsing remitting MS; SPMS, secondary progressive MS; IFN, interferons; CIS, clinically isolated syndrome; MHC, major histocompatibility complex; BBB, blood brain-barrier; Th, T helper; TGFβ, the transforming growth factor beta; IL, interleukin; TNFα, tumour necrosis factor; IFNγ, interferons gamma; VLA4, very late activation protein 4 receptor, alpha 4 subunit; VCAM-1,vascular cell adhesion molecule; CNS, central nervous system; PML, progressive multifocal leukoencephalopathy; Nrf2, nuclear factor erythroid 2–related factor 2; ADCC, antibody-dependent cellular cytotoxicity; NK, natural killer cell.
*The drug was withdrawn from the market in March 2018 following reports of serious and potentially fatal immune reactions affecting the brain (including encephalitis and meningoencephalitis), liver, and other organs.
[a] (Kappos et al., 2018), [b] (Olsson et al., 2014), [c] (Sorensen et al., 2014).

All mentioned agents in Table 1.2 act by modulating and/or suppressing the immune system at different levels and with different mechanisms of action, which may make them more effective. However, the efficacy, tolerability and safety profile vary greatly across therapies for MS, ranging from combinations of modest safety and effect to options that are highly effective but at increased risk of serious adverse events, which may be fatal in rare cases (Gajofatto and Benedetti, 2015). For example, interferon- β and glatiramer acetate treatment need frequent subcutaneous or intramuscular injections and are only moderately effective, but very rarely have life-threatening adverse effects, while teriflunomide and dimethyl fumarate are administered orally and have equal or better efficacy, but have more potentially severe adverse effects (Soelberg Sorensen, 2017). Fingolimod, natalizumab and alemtuzumab are highly effective therapies, but have more serious adverse effects, some of which may be life-threatening (Soelberg Sorensen, 2017). In addition, the most available DMTs have a favourable impact on RRMS, while there are only a few available DMTs for the other forms of MS (e.g., Siponimod in SPMS and Ocrelizumab in PPMS) (Thompson A, 2018) . Therefore, there is an unmet clinical deemed for developing highly effective therapies with a better safety profile, which is even more significant when considering preventive drugs for individuals at high risk of disease (Jacobs et

al., 2020). One response to this situation is drug repurposing, a strategy to speed up the traditional process of drug discovery by identifying a novel clinical use for drugs that have already proven to be safe and effective in humans and are approved for other indications (Sultana et al., 2020). In the third and fourth chapters, I describe drug repurposing opportunities for MS.

1.10. Genome-wide association studies in multiple sclerosis

Genome-wide association studies (GWAS) have made major progress in discovering the risk loci associated with MS in the last decade. These loci were found to be involved in a wide range of biological pathways related to the risk of developing MS. Most importantly, GWAS studies were able to identify MS loci outside the MHC for the first time. Several GWASs and meta-analyses were conducted by the International Multiple Sclerosis Consortium (IMSGC). The most informative studies include the GWAS in collaboration with the Welcome Trust Case Control Consortium 2 (WTCCC2), which involved 13,990 cases and 24,672 controls (discovery plus replication) of European descent (Sawcer et al., 2011). In this study, a total of 52 non-MHC loci were identified, in which 23 previously reported risk loci were replicated, and a further 29 novel risk loci were identified as genome-wide significant (Sawcer et al., 2011). Most of these loci were found nearby genes encoding for immune-related proteins, and the other loci were previously associated with other autoimmune diseases (Didonna and Oksenberg, 2015).

Two years later, the follow-up to the 2011 GWAS was conducted using the ImmunoChip custom genotyping array and involving 29,300 subjects with MS and 50,794 healthy controls (discovery plus replication) of European ancestry (Beecham et al., 2013a). In this landmark study, a total of 97 genome-wide significant loci outside the MHC region were identified; of these loci, 48 were new and 49 corresponded to previously identified susceptibility loci for MS (Beecham et al., 2013a).

In 2019, the IMSGC conducted the largest meta-analysis GWAS in MS, which consisted of 47,429 MS subjects and 68,374 healthy control subjects (discovery plus two independent replications of data)

of European ancestry (Consortium*† et al., 2019). Remarkably, this study was able to identify 233 statistically independent associations with MS susceptibility that are genome-wide significant, in which MHC contains 32 of these associations (Consortium*† et al., 2019) (more details in methodology section 2.2.3). Altogether, these findings enhance the catalogue of MS risk loci and enrich our knowledge of the immune system processes implicated in MS.

Surprisingly, GWAS studies on MS severity have received relatively less attention. So far, the IMSGC has conducted only three GWAS to identify loci influencing disease severity (as measured in terms of MS severity score). The first GWAS study comprised 1470 MS cases (Briggs et al., 2011). The second study consisted of 7,069 MS cases obtained from the 2011 GWAS (Sawcer et al., 2011), while the third study employed 7,934 MS cases obtained from the discovery phase of the ImmunoChip GWAS (Beecham et al., 2013a). However, no single loci achieved genome-wide significance, indicating that MS susceptibility loci are unlikely to influence MS severity. This finding highlights the need for further larger studies with a primary focus on disease severity to enhance the power to identify loci that affect MS severity, which is essential to furthering our understanding of disease mechanisms and, most importantly, to guide the development of effective therapeutic approaches.

1.11. Mendelian randomisation approach

Epidemiological studies, such as observational and randomised controlled trials (RCTs), are the most straightforward studies to identify determinants (causes, risk factors) of health-related states and events. Although RCTs are considered the gold standard study design for inferring causality, RCTs can be limited by time (i.e. they are unlikely to be conducted in the short term) and RCTs make it unfeasible to randomise people to every risk factor due to ethical problems (Howell et al., 2018); for example, pregnant women cannot be randomised to drink wine during pregnancy. On the other hand, observational studies can be subject to biases from confounding, measurement error and reverse causation, leading to spurious associations, which can be difficult to eliminate even through statistical

adjustment (Howell et al., 2018). Additionally, observational studies cannot distinguish correlation from a causation relationship (Davey Smith and Hemani, 2014).

One useful method to appraise causality within observational epidemiology, which is relatively quicker and easier than RCT studies and overcomes some of the limitations inherent in conventional epidemiologic studies, is Mendelian randomisation (MR) (Williams et al., 2020).

1.11.1. Mendelian randomisation for causal inference

In simple terms, MR is a type of “instrumental variable” analysis that uses genetic variants, such as SNPs, robustly associated with exposures as proxies for the risk factors of interest to investigate their causal effect roles on outcomes (Bennett and Holmes, 2017, Burgess et al., 2017a, Howell et al., 2018). MR takes advantage of Mendel’s laws of inheritance, in which the germline genetic variants are randomly distributed during meiosis and independent of non-genetic factors, including environmental risk factors, confounding factors or disease processes (De Silva et al., 2011, Williams et al., 2020). Thus, using such genetic variants in MR analysis limits the bias associated with reverse causation and confounding in conventional epidemiologic studies. The fundamental principle utilised in this approach is as follows: if a genetic variant either alters the level of, or mimics the biological effects of, a modifiable biomarker that is causal in disease, then these genetic variants should also be related to disease risk to the extent predicted by the effect of the genetic variant on exposure to the biomarker (Smith et al., 2008, Bennett and Holmes, 2017). For example, nine SNPs were associated with lower low-density lipoprotein cholesterol (LDL-C) (Ference et al., 2012). Using all these SNPs in MR was associated with a highly consistent reduction in the risk of coronary heart disease (CHD) per unit lower LDL-C (Ference et al., 2012) (**Figure 1.6**).

1.11.2. Mendelian randomisation for drug repurposing

The initial interest in MR was mainly focused on elucidating the causal effect of environmental exposures on medically relevant outcomes (Zheng et al., 2017). Since genotypes are assigned almost independently of environment when they are inherited from parents, MR can be described as ‘nature’s randomised controlled experiment’ in which germline genetic variants are regarded as randomised proxies for exposure of interest, in the same way that the allocation group in an RCT is a proxy for an intervention of interest (Howell et al., 2018, Sonehara and Okada, 2021) (**Figure 1.6**). MR offers virtual RCT opportunities without actual intervention (Sonehara and Okada, 2021). Therefore, a significant amount of attention has been directed to MR across a wide range of domains, including the development of pharmaceutical agents (i.e., drug target identification, repurposing, validation and side effect identification) (Zheng et al., 2017). This approach has been significantly potentiated by the comprehensive coverage of GWAS that has been made publicly available and the increasing availability of high-throughput genomic and proteomic technologies (Zheng et al., 2017, Acosta et al., 2021).

Drug repurposing MR studies use genetic variants acting in ‘cis’ on druggable protein levels or gene expressions that encode druggable proteins (i.e. proteins are known to be targeted by drugs) as a proxy for exposure to drugs (e.g. statins) to predict their causal association with an outcome (Gaziano et al., 2021). Once the casual association is demonstrated, this result can be used to inform the potential repurposing of the drug to treat the disease.

The distinction between conventional MR and drug target MR is that conventional MR establishes causal relationships between biomarkers or traits and an outcome, while drug target MR aims to address whether modifications of a specific drug target or protein will have an effect on the outcome (Acosta et al., 2021).

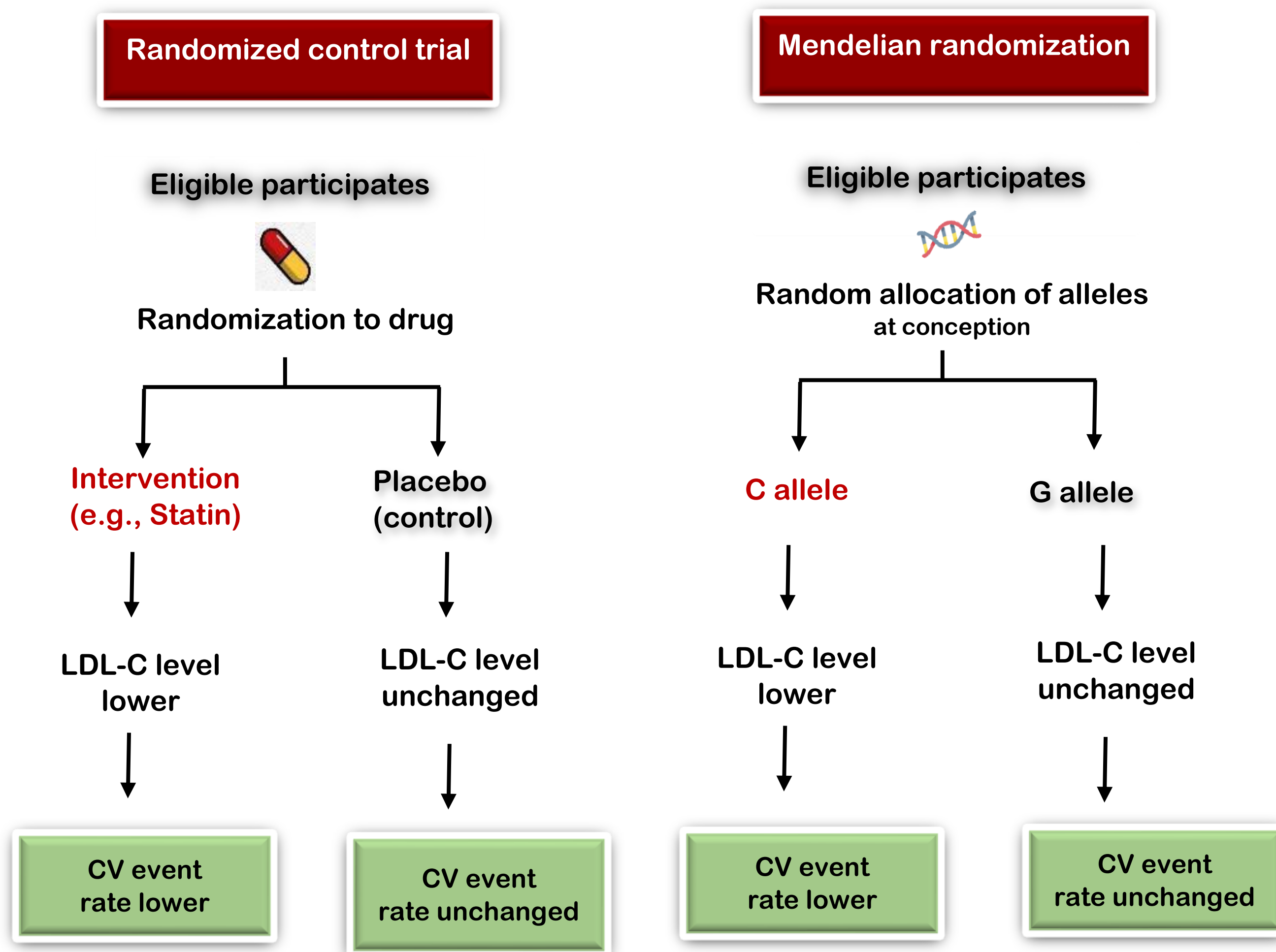


Figure 1.6: Schematic comparison between a randomised control trial (RCT) and Mendelian randomised (MR) study using LDL-C (low-density lipoprotein cholesterol), statin and CV (cardiovascular disease) as an example. The left-hand panel of the figure represents the RCT in which the participants are randomised into the intervention group (who received statin) and placebo (control group). In the intervention group, statin reduces LDL levels and prevent CV events, while there is no change in the LDL-C levels and CV event among placebos. The right-hand panel of the figure represents the MR in which the participants are randomly assigned to an exposed group who carries allele **C**, which represents the intervention arm in RCTs, and a control group who carries allele **G** (normal control alleles), which represents the placebo arm in RCTs. Allele C mimics lifelong exposure to lower levels of LDL-C. Thus, the carriers of allele C have a lower CV event, which mimics the statin intervention in RCT. This figure is adapted from (Bennett and Holmes, 2017).

1.12. Mendelian randomisation in multiple sclerosis

Over the past several years, a massive investment in large-scale GWASs has resulted in discovering reliable genetic variants for a wide range of phenotypes, including modifiable environmental exposures, such as circulating vitamin levels, as well as complex human behaviours, including nicotine dependence (Jiang et al., 2018). These efforts provide an unprecedented opportunity for genetic epidemiology in particular by utilising the MR design (Jiang et al., 2018). The success of the MR approach in uncovering causal relationships between several environmentally modifiable exposures and the risk of developing MS has been demonstrated in numerous works. For example, MR showed that genetically lowered vitamin D levels and genetically elevated BMI were strongly associated with an increased risk of MS development (Mokry et al., 2016, Rhead et al., 2016). This finding is in agreement with observational studies that observed a correlation between vitamin D level, BMI and MS risk (Mokry et al., 2016, Rhead et al., 2016). Further, MR found evidence of an association between higher age at puberty and decreased risk of MS, which is in line with previous epidemiologic studies (Harroud et al., 2019).

On the other hand, MR reported conflicting results in the observational findings. For example, MR found no evidence for the causal role of coffee consumption on MS risk (Lu et al., 2020). This finding does not support observational studies that repeatedly report an association between coffee consumption and increased risk for MS (Lu et al., 2020). Further, MR does not support a causal effect of genetically determined serum uric acid levels and adiponectin on the risk of MS, and these MR results also do not support previous studies that suggested a role for these molecules in the development of MS (Niu et al., 2020, Devorak et al., 2017). These conflicting results between MR and traditional observational epidemiology are possibly due to differences in the study population, small sample size, confounding, misclassification or reverse causality (Lu et al., 2020).

The above studies provide valuable insights into the aetiology of MS and highlight the advances that have been made in identifying causality in observational research. In Chapters Three, Five and Six, I sought to continue uncovering the causality between other risk factors that have not yet been studied or have been studied with limitations, with a primary focus on MS severity.

1.13. Thesis motivation and aim

Observational studies have reported associations between several risk factors and increased MS risk, as mentioned above. However, it is not clear whether these associations are causal. Accurate assessment of environmental risk factors is crucial not only to understand MS risk aetiology but also to improve prevention strategies, identify novel therapeutic targets and delay disease progression among MS patients. Therefore, taking advantage of the availability of large-scale GWASs, for my doctoral work, I aimed to apply MR to uncover causal relationships between several environmentally modifiable exposures and the risk of developing MS. I believe this would enrich our knowledge of MS aetiology and may carry meaningful implications for MS patients' care by aiding clinical diagnosis and perhaps treatment.

Although much attention has been directed to applying MR design to MS risk, MS severity has received less attention, perhaps because the MS severity GWASs are not publicly available. Therefore, I aimed to shed light on MS severity by requesting access to the GWAS datasets and then taking forward this data and applying an MR design to assess the causal role of several risk factors in the risk of worsening MS severity. To my knowledge, my work here is the first to apply MR to MS severity.

1.13.1. Specific objectives of this thesis

Numerous epidemiological studies have reported an association between lipid and disability progression, but the role of lipids in MS risk development lacks attention. I used the MR approach to dissect the causal nature of the association between lipid, MS risk and severity (**Chapter 3**).

Recently, statins have received much attention due to their beneficial pleiotropic effects. Several MR studies have been performed to explore the repositioning opportunities and side effects of statins. The findings from these studies suggest that statins may reduce the risk of developing health-related phenotypes, such as cancer. Additionally, the phase 2 trials of statins on SPMS showed promising results. These findings highly motivated me to perform MR analysis in an effort to examine whether statin can also be used to prevent MS risk and to delay MS severity (**Chapter 3**).

Taking advantage of public expression quantitative trait locus datasets, I aimed to apply MR to the druggable genome in an effort to identify new drug-targeting mechanisms for MS with genetic support—more importantly, for MS severity (**Chapter 4**).

Recent studies using MR suggest that the relationship between obesity and the development of MS is causal. These studies used loci associated with body mass index (BMI) as a measure of obesity. It has been shown that BMI and other anthropometric measures, such as height and non-fat, share common loci. Thus, the reported obesity-MS risk association may be driven by other measures. Therefore, I sought to perform MR analysis considering 21 anthropometric measures to distinguish between the contributions of fat and non-fat to the obesity-MS risk association and to assess the role of these measures on MS severity (**Chapter 5**).

There is a considerable body of literature linking MS and stroke. Recent MR found no effect of genetic liability to MS on the risk of stroke. However, this study did not examine the causal effect of stroke on MS risk. Therefore, I aimed to conduct MR to investigate the potential for a causal effect of stroke on MS risk and severity (**Chapter 6**).

Chapter 2 Method

2.1. Overview

MR analysis was conducted in this thesis using a two-sample MR approach. This approach utilises genetic variants, specifically SNP, to estimate the casual effect of an exposure, such as a biomarker, on a disease outcome in a non-experimental (observational) setting (Burgess et al., 2015). The two-sample MR approach uses only summary statistics in which the SNP-exposure associations and SNP-outcome associations are estimated in different GWASs but the same underlying population (Burgess et al., 2015). To conduct a valid MR study, genetic variants should satisfy three key assumptions (**Figure 2.1**):

1. The genetic variant should be strongly (p-value $\leq 5 \times 10^{-8}$) associated with the risk factor of interest (the relevance assumption) (Burgess et al., 2015).
2. The genetic variant should not be associated with confounders of the risk factor–outcome association (the independence assumption) (Burgess et al., 2015).
3. The genetic variant should only affect the outcome via the risk factor of interest, not via other casual pathways, either directly to the outcome or through confounders (Burgess et al., 2015, Zheng et al., 2017). This assumption is referred to as the exclusion restriction criterion or no pleiotropy (Burgess et al., 2015, Zheng et al., 2017).

Violation of any of these three assumptions could lead to unreliable causal estimates. These assumptions, however, cannot all be tested empirically (Burgess et al., 2015). Only the first assumption can be tested to assess whether the association between genetic variants and risk factors is strong or weak by using the mean F-statistic (Bowden and Holmes, 2019, Bowden et al., 2016). Meanwhile, the second and third assumptions cannot be tested empirically because the unmeasured or unknown confounders with genetic variants cannot be assessed.

Therefore, I conducted several sensitivity analyses to detect and adjust for pleiotropy to ensure the reliability of the estimates (more details are provided below).

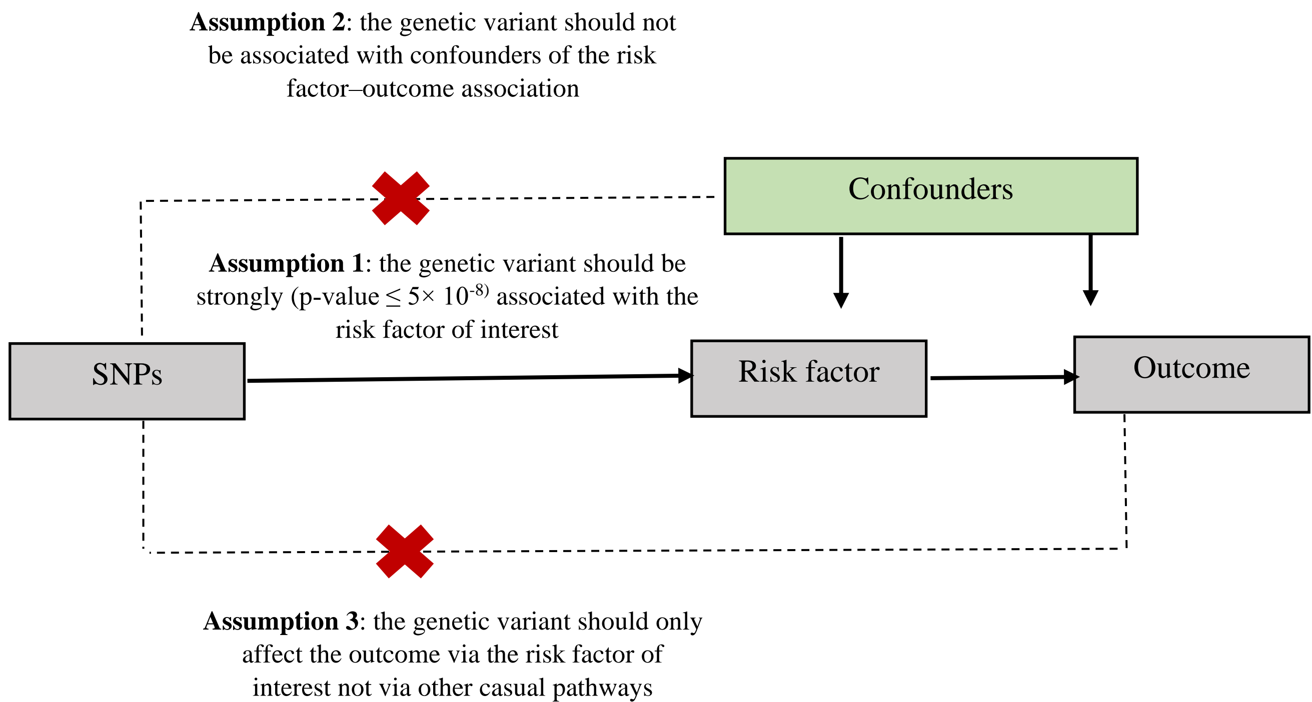


Figure 2.1: Schematic diagram illustrating the three key assumptions for MR.

A two-sample MR workflow in this thesis can be summarised into four main stages (**Figure 2.2**):

1. Dataset identification and summary statistics extraction.
2. Prioritising and validating the genetic instruments.
3. Estimating the causal effect.
4. Additional analyses (multivariable MR and bidirectional MR).

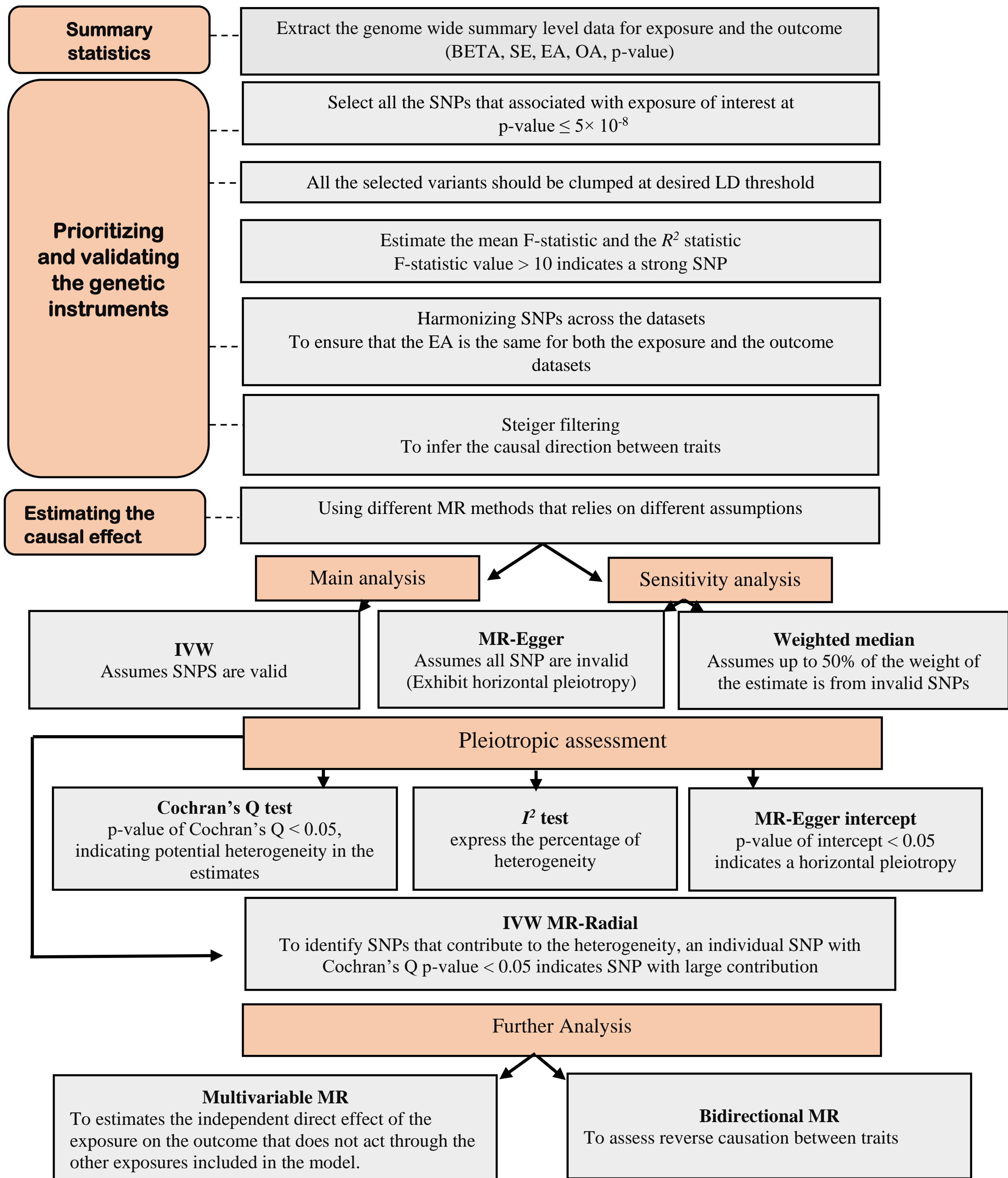


Figure 2.2: Schematic diagram outlining the standard analysis workflow for an MR study in general and for this thesis specifically. Effect size stands for Beta / log odds ratio. Abbreviation: SE, standard error; EA, effect allele; OA, other alleles; LD, linkage disequilibrium; SNP, single nucleotide polymorphism; IVW, inverse-variance weighted; MR, mendelian randomization.

2.2. Dataset identification and summary statistics extraction

2.2.1. GWAS datasets descriptions overview

This thesis uses genome-wide summary-level data from eight cohorts (**Table 2.1**) divided into two groups: the first group of datasets included GWAS data for exposures, such as the Global Lipids Genetics Consortium (GLGC), eQTLGen Consortium, the Genotype-Tissue Expression (GTEx) Consortium and the Neale lab consortium. The second group of datasets included the outcome GWAS datasets, MS risk GWAS and MS severity GWAS. All these datasets are publicly available and can be downloaded from the relevant website in **Table 2.1**, except the MS datasets were available upon request; I have requested MS risk-GWAS from the IMSGC and MS severity-GWAS directly from the corresponding author, Professor Jacob McCauley. In this section, I will briefly describe each dataset and the purpose of using it in turn. Together, these data provide genetic data on a large sample size, which offers opportunities for performing MR analyses. Such large sample sizes are required in MR to detect the small effect sizes that are common when using genetic data to study complex phenotypes [51]. For each dataset, I extracted the summary statistics of GWAS for each SNP, included β-coefficients/log odds ratio (effect size), standard errors (SE) for the effect size, the effect alleles ('EA'; the allele for which the effect size is calculated), non-effect alleles ('OA'; the complementary allele of the effect allele), effect allele frequency (the relative frequency of the effect allele from 1000 G EUR sample) and p-values [52]. All the datasets that I have used in this thesis were mapped to the human reference genome GRCh37.

Table 2.1: Summary of GWAS used in the thesis.

Trait (consortium)	Sample size	Date	Website	Reference
Lipid fractions (GLGC)	188,578	2013	http://csg.sph.umich.edu/willer/public/lipids2013/	(Willer et al., 2013)
whole blood *cis*-eQTL (eQTLgen consortium)	31,684	2021	https://www.eqtlgen.org/cis-eqtls.html	(Võsa et al., 2021)
whole blood *cis*-seQTL (GTx consortium)	670	2020	https://www.gtexportal.org/home/datasets	(Consortium, 2020)
Anthropometric-related measures (Neale Lab UK Biobank GWAS)	361,194	2018	https://docs.google.com/spreadsheets/d/1kvPoupSzsSFBNSztMzl04xMoSC3Kcx3CrjVf4yBmESU/edit#gid=178908679	http://www.nealelab.is/uk-biobank
HDL-C (Neale Lab UK Biobank GWAS)				
MS risk	115,803	2019	https://imsgc.net/?page_id=31	(Consortium*† et al., 2019)
MS severity	7,069	2011		(Sawcer et al., 2011)
Brain *cis*-eQTL (PsychENCODE)	1,387	2018	http://resource.psychencode.org/#	(Wang et al., 2018)
Stroke and its subtypes (METASTROKE)	521,612		https://www.ebi.ac.uk/gwas/publications/29531354	(Malik et al., 2018)

2.2.2. Exposure datasets

2.2.2.1. Lipid fractions data

The summary statistics data for the genetic instruments associated with blood lipid fractions (low-density lipoprotein cholesterol (LDL-C), triglycerides (TG) and high-density lipoprotein cholesterol (HDL-C)) were taken from GLGC GWAS to investigate the association between lipids and MS (Willer et al., 2013). This GWAS study examined individuals of European ancestry, including 188,578 participants from 60 studies. Individuals known to be on lipid-lowering medications were excluded from the study (Willer et al., 2013).

This GWAS study identified 157 loci associated with blood lipid fractions that achieved genome-wide significance (p-value $< 5 \times 10^{-8}$). These loci were at a distance of one megabase (Mb) pair intervals and were nearly independent (pairwise r^2 value < 0.10) (Willer et al., 2013). The effect estimates for each lipid fraction were in standard deviation (SD) units and were estimated using inverse normal transformed residuals of lipid levels after adjusting for age and sex (Willer et al., 2013). The MR results obtained with the GLGC datasets were replicated using the summary statistics data for lipid fractions from the Neale Lab consortium; refer to Section 2.2.2.3 for more details about the Neale Lab consortium.

2.2.2.2. Expression quantitative trait loci data

To examine the causal links between the expression levels of a set of genes and MS, I obtained blood and brain *cis*-expression quantitative trait loci (*cis*-eQTL) from publicly available data, including eQTLGen, GTEx and PsychENCODE consortia. I used only *cis*-acting SNPs because they are regarded as a robust instrument and more likely to satisfy the key assumptions for MR due to their direct effects on the expression level of the proximal genes, in contrast to the trans-acting SNPs that affect the expression of distant genes and hence these SNPs are more prone to violate MR assumptions (Porcu et al., 2019).

The eQTLGen consortium performed a large-scale meta-analysis of up to 31,684 blood samples of individuals of European ancestry from 37 cohorts (Võsa et al., 2021). The *cis*-eQTL data is available for SNP located within less than one megabase (Mb) around each of 16,987 genes in the eQTLGen consortium (Võsa et al., 2021).

The GTEx project (version 8) measured gene expression in the whole blood of 670 individuals, where most of the donors were of European ancestry (Consortium, 2020). In this project, the *cis*-eQTL data is available for SNP located within two Mb window around the transcription start site (Consortium, 2020).

The PsychENCODE consortium performed a meta-analysis of eQTL studies of brain prefrontal cortex samples from1,387 individuals of mostly European ancestry (Wang et al., 2018). Only *cis*-eQTL data for SNPs located within one Mb window around each gene is available in the PsychENCODE consortium (Wang et al., 2018).

2.2.2.3. Neale Lab UK Biobank GWAS

To investigate the association between anthropometric measures and MS, summary statistics data for anthropometric-related measures were obtained from the Neale Lab consortium. This consortium was conducted on 4,236 phenotypes with a sample size of 361,194 persons of white-British ancestry from the UK Biobank. The UK Biobank is a population-based cohort that recruited over 500,000 participants aged 40–69 years in 2006–2010 to provide biological samples and comprehensive clinical information (Sudlow et al., 2015). The UK Biobank was established to identify the genetic and nongenetic determinants of human diseases in middle to older aged individuals (Sudlow et al., 2015).

For anthropometric-related measures, summary statistics data were obtained for the genetic variants associated with 21 anthropometric-related measures. For the purpose of analysis, I divided these measures into two categories: the adiposity-related measures included BMI, weight, fat mass (FM) and fat percentage (FP) for the whole body, the upper limbs (right arm and left arm), lower limbs (right leg

and left leg) and trunk. The second category included height and non-fat mass (NFM) for the whole body, upper limbs, lower limbs and trunk.

The anthropometric-related measures were collected from UK Biobank participants by trained staff at a baseline assessment centre visit. The participants' heights were obtained using the Seca 240 cm height measure. BMI was estimated as weight in kilograms (kg) divided by height in metres squared (m^2). The participants' weight and body composition data were collected using a Tanita BC-418 MA body composition analyser (Tanita, Tokyo, Japan). This device measures bioelectrical impedance in the body and produces a print-out of segmental readings of fat percentage, fat mass and non-fat mass for the whole body, limbs and trunk.

The association summary statistics between the genetic variants and the anthropometric measures were generated using a linear regression model that included the first 20 principal components, age, sex, age2, sex∗age and sex∗age2 as covariates to adjust for both sexes (Howrigan, 2019).

Further, to generate a heatmap plot, I downloaded the genetic correlation data between anthropometric measures from the Neale Lab (http://www.nealelab.is/blog/2019/10/10/genetic-correlation-results-for-heritable-phenotypes-in-the-uk-biobank). These genetic correlation datasets are made publicly available, and they have been estimated through cross-trait linkage disequilibrium score regression using summary statistics data (McInnes et al., 2018).

2.2.2.4. GWAS dataset for stroke

Summary statistics for SNPs associated with stroke and its subtypes were derived from METASTROKE consortium, a large-scale multi-ancestry stroke GWAS meta-analysis study, including 521,612 subjects in total (67,162 stroke cases and 454,450 normal controls) (Malik et al., 2018). I obtained the summary dataset for European participants only, which includes 34,217 cases and 406,111 controls for ischemic stroke (IS), 5,386 cases and 192,662 controls for small vessel strokes (SVS), 4,373 cases and 40,6111 controls for large-artery atherosclerotic stroke (LAS) and 7,193 cases

and 406,111 controls for cardioembolic stroke (CES). The phenotypic variance explained in IS, SVS, LAS, and CES were 0.9%, 0.8%, 1.7% and 1.4%, respectively.

2.2.3. MS outcome data

The most recent cohort of the IMSGC meta-analysis jointly analysed (one discovery cohort and two independent replication datasets) data from 47,429 cases with MS and 68,374 healthy controls of European descent to provide a comprehensive genetic evaluation of MS susceptibility (Consortium*† et al., 2019). In this GWAS, 233 genetic variants were identified with strong evidence (p-value $< 5 \times 10^{-8}$) of association with MS susceptibility and explained about 39% of the genetic predisposition to MS (Consortium*† et al., 2019). Of the 233 variants, 200 autosomal susceptibility variants are located outside of the MHC region and account for almost 19% of the MS heritability (Consortium*† et al., 2019). In this thesis, I used the summary statistics data from the discovery cohorts, which included 14,802 cases with MS and 26,703 healthy individuals as GWAS for MS risk. Due to complex linkage disequilibrium structures and a high potential for pleiotropy in the MHC region, 12 Mbps around this region (from 24 to 35 megabase pairs of chromosome 6; GRCh37) were excluded from MS discovery GWAS. For the purposes of MR bidirectional analysis (also known as reverse causation analysis), where MS risk is used as an exposure, I used summary statistics data for 200 autosomal susceptibility variants. Refer to Section 2.3.2.4.2 for more details about MR bidirectional analysis.

For MS severity, I obtained the summary statistics data from Professor Jacob McCauley, the corresponding author of the original publication (Sawcer et al., 2011). The MS severity data has been generated from a genome-wide scan performed in MS cases (7,069 cases) to identify genetic variants that might influence MS severity (Sawcer et al., 2011). In this severity-based analysis, the MS cases obtained from the discovery phase of the primary analysis of susceptibility of the 2011 GWAS, which included 9,772 cases and 17,376 controls (Sawcer et al., 2011). Of the 9,772 cases, the disease severity (as measured by MS severity score) was available for only 7,069 cases, thus association analyses were only performed on 7,069 cases (Sawcer et al., 2011). No genetic variants with strong evidence (p-value

$< 5 \times 10^{-8}$) for association with MS severity were identified in that severity-based analysis (Sawcer et al., 2011). Therefore, I was not able to perform a bidirectional MR analysis using this data.

2.3. MR analysis

2.3.1. Prioritising and validating the assessment of genetic variants

In the MR approach, it is essential to ensure that the genetic instruments are robustly associated with the exposure of interest. Therefore, the selection process for the genetic instruments in this thesis followed the following steps:

2.3.1.1. Identify the genetic instruments

SNPs associated with exposure of interest at a GWAS threshold of statistical significance (p-value $\leq 5 \times 10^{-8}$) were selected. Using such a stringent threshold in MR studies helps avoid a false positive association between SNPs and the exposure of interest in GWAS and avoids weak instrument bias (Wang et al., 2021).

2.3.1.2. Independence assessment

The selected genetic instruments were clumped to generate an independent genetic instrument (SNP) because variants in linkage disequilibrium (LD) with other nearby variants may influence the outcome via other exposures rather than the exposure of interest. If that were the case, LD might violate the exclusion restriction MR assumption and bias MR estimates (VanderWeele et al., 2014). Thus, it is necessary to ensure that the selected genetic instruments for an exposure are independent unless measures are taken in the MR analysis to account for any correlation structures that arise through LD (Hemani et al., 2018b).

To do so, I used the *clump_data* function in the "TwoSampleMR" R package, which uses European ancestry in the 1,000 genomes project as a reference panel (Hemani et al., 2018b). The clumping criterion was set to retain genetic instruments with the smallest p-values above the selected clumping threshold within a 10,000 (Kb) window (Hemani et al., 2018b). In this thesis, I used two clumping

thresholds: a stringent ($r^2 < 0.001/0.01$) and a liberal ($r^2 < 0.2/0.4$) clumping threshold. More details about how and which clumping threshold was used will be provided in the results chapters.

2.3.1.3. Strength of the genetic instrument assessment

Next, to judge the strength of the association of instrument and exposure, I used the mean F-statistic and the R^2 statistic, where higher values for both of these statistics indicate the suitability of genetic instruments for MR analysis. The F-statistic was calculated for each genetic instrument included in the analysis, and then the mean was obtained (Bowden and Holmes, 2019, Bowden et al., 2016). The mean F-statistic for independent (i.e. uncorrelated) genetic instruments can be estimated as follows (Bowden and Holmes, 2019, Bowden et al., 2016) :

$$mean\ F = \frac{1}{L}\sum_{1=j}^{L}\frac{\gamma^\wedge 2_j}{\sigma 2_{xj}}$$

where j = SNP, $\gamma^\wedge 2_j$ = the squared effect estimates of the SNP-exposure, $\sigma 2_{xj}$ = the squared standard error of the SNP-exposure and L = the total number of SNPs. A mean F value of greater than 10 has been proposed to indicate a strong instrument (i.e. strongly predict the exposure) as well as to avoid a weak instrument, that is, the genetic variants that explain a small proportion of the overall variation in a given exposure (Sheehan and Didelez, 2011, Burgess et al., 2011). The effect of such weak instruments is shown to bias the inverse variance weighted (IVW) estimate towards the null (Bowden et al., 2016, Bowden and Holmes, 2019).

The other important measure is the R^2 statistic, which is the proportion of variance (R^2) in the risk factor explained by the genetic variants included in the analysis. If the coefficient for the association of an SNP with the risk factor is given in standard deviation (SD) units (risk factor is continuous) (Burgess et al., 2016), R^2 can be estimated from the given formula:(Au Yeung et al., 2018)

$$R^2 = \sum_{j=1}^{k} R_j^2$$

where R_j^2 = 2 ×EAF × (1 − EAF) ×Beta^2× Var, EAF = effect allele frequency, Beta = beta coefficient for the genetic instruments, and var = the variance of the risk factor, which is equal to one because the beta coefficient refers to a change in 1 standard deviation (Au Yeung et al., 2018).

2.3.1.4. Harmonising genetic instruments across the dataset

Once the genetic instruments for the exposure of interest were selected, the effect estimates (i.e. summary statistics) for each genetic instrument on the outcome GWAS were obtained. As a next step, I carried out the harmonisation by using the *harmonise_data* function in the "TwoSampleMR" R package to ensure that the effect estimate of a given genetic instrument on the exposure and the effect estimate of the same instrument on the outcome dataset correspond to the same effect allele (Hemani et al., 2018b). Palindromic SNPs and SNPs with incompatible alleles were discarded from the analysis.

SNPs whose alleles on the forward strand are similar to those on the reverse strand (C/G on forward is G/C on the reverse, or T/A on forward is A/T on the reverse) are called palindromic SNPs, which can lead to ambiguous effect alleles in the summary datasets for the exposure and outcome.

Incompatible alleles are a situation in which the alleles of SNP-exposure and SNP-outcome are not similar. For example, a given SNP with T/C alleles for the exposure and T/G alleles for the outcome (Hemani et al., 2018b). This occurs due to either an error in the data or the use of different builds.

Thus, flipping will not resolve these differences among alleles, thus SNPs with such ambiguity are excluded from the analysis (Hemani et al., 2018b).

2.3.1.5. Directionality assessment

MR analysis assumes that the genetic instruments used to proxy the exposure of interest exert their primary association on that exposure and that any association with the outcome is a result of a causal effect of the exposure of interest on the outcome (Zheng et al., 2020, Zheng et al., 2019). However, this is not always the case, and there is a possibility that some genetic instruments are primarily associated with the outcome (Zheng et al., 2020, Zheng et al., 2019). Using such variants as genetic instruments to proxy the exposure of interest in MR analysis could, then, be misleading. This is because any existing causal relationship between the exposure and the outcome is, in reality, due to the causal effect of the outcome on the exposure of interest (Zheng et al., 2020). For this reason, at the final step in the selection and validity, I used the *steiger_filtering* function in the "TwoSampleMR" R package to restrict genetic instruments to those that have the primary strongest effects on the exposure to ensure the causal association is in one direction (i.e. exposure to outcome) and not in the opposite direction (i.e. outcome to exposure) (Zheng et al., 2020, Zheng et al., 2019, Hemani et al., 2018b).

The *steiger_filtering* function infers the causal direction between phenotypes by flagging the genetic instrument with 'TRUE' if the instrument explains more of the exposure variances than the outcome variances, which indicates causality in the expected direction, i.e. the exposure influences the outcome (Zheng et al., 2020). If a genetic instrument explains more of the outcome variances than the exposure variances, this instrument will be flagged as 'FALSE', which indicates causality in the reverse direction, i.e. the outcome influences the exposure (Zheng et al., 2020). In each MR analysis, I kept only variants that were 'TRUE'.

2.3.2. Estimating the causal effect

2.3.2.1. Overview

In this section, the statistical analysis was carried out using different MR methods to estimate the causal effect of the risk factor of interest on MS. I used the IVW method as the primary analysis to estimate the causal effects. I then carried out several sensitivity analyses to assess the reliability of the estimates obtained from IVW. Briefly, sensitivity analyses included MR-Egger, weighted median and MR-radial. In addition, as a part of sensitivity analyses, I also used Cochran's Q, the I^2 statistic and the MR-Egger intercept to assess the heterogeneity and pleiotropy, respectively, across the genetic variants. Further, multivariable MR and bidirectional MR analyses were used to obtain the direct effect of an exposure on MS and to examine MS's influences on a range of risk factors, respectively. Table 2.2 shows the methods used to estimate and evaluate the causal effect estimate. Figure 2.3 shows the interpretation scheme followed in this thesis to interpretation the results.

Table 2.2: The methods used to estimate and evaluate the causal effect estimate

	Method	Purpose
MR estimate test	IVW	To estimate the causal effect of an exposure on an outcome
Sensitivity tests	MR-Egger	To assess the reliability of the IVW estimate
	Weighted median	To assess the reliability of the IVW estimate
Pleiotropic assessment	MR-Egger intercept	To detect pleiotropy
	Cochran's Q and I^2 tests	To detect heterogeneity
	MR-Radial	To detect invalid variants
Further Analysis	Multivariable MR analysis	To estimate the direct causal effect of an exposure of interest on the outcome that does not act through the other exposures included in the model.
	Bidirectional MR	To explore whether the liability to the outcome would exert a change in the exposure.

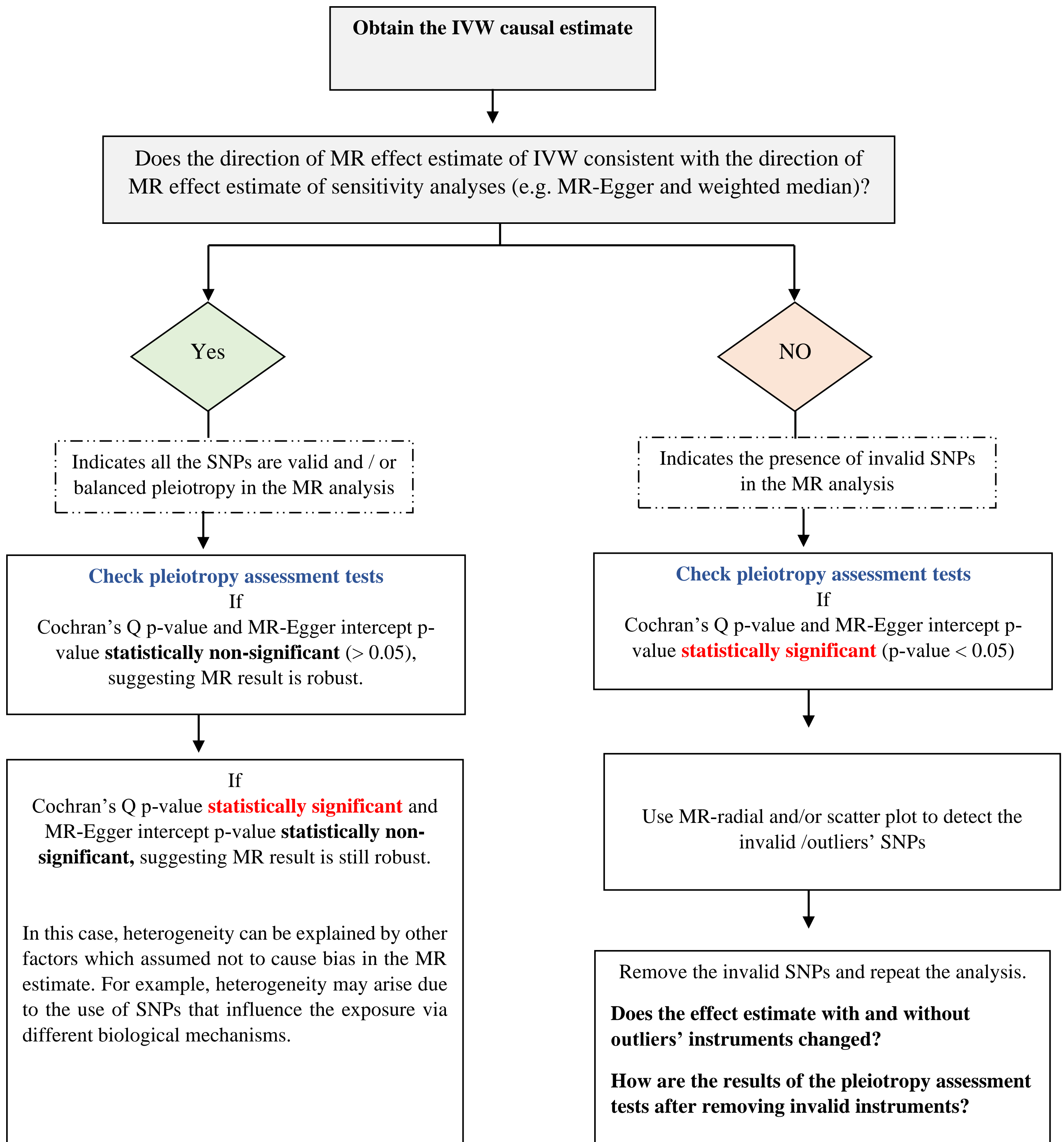


Figure 2.3: A flowchart illustrating how the MR results have been interpreted in the present thesis. Abbreviations: MR, Mendelian randomization; SNP, single nucleotide polymorphisms; IVW, inverse-variance weighted.

Note: Caution should be taken when removing the SNPs from the analysis, as this strategy would reduce the SE and may lead to over-fitting.

2.3.2.2. Inverse-variance weighted

IVW is a straightforward way to obtain unbiased causal effect estimates from multiple SNPs. In the IVW analysis, the causal estimate of the exposure on the outcome for each SNP is first obtained separately using the Wald ratio (or ratio estimate) approach. The Wald ratio estimate is obtained by dividing the SNP-outcome association estimate by the SNP-exposure association estimate (Burgess et al., 2017b).

The Wald ratio estimates from each SNP were then pooled in a multiplicative random-effects model and weighted using the inverse standard error (SE) of the SNP-outcome association estimates as weightings to obtain the average casual effect, indicating the increase in outcome per unit change in the exposure (Burgess et al., 2013, Burgess and Bowden, 2015). The purpose of using the inverse variance of the SNP-outcome association as weight is to allow more robust SNPs to make a larger contribution to the causal estimate (Hemani et al., 2018b).

The IVW can be performed using a fixed or random-effects model. I preferred the multiplicative random-effects model to the fixed-effect model because it accounts for heterogeneity between Wald estimates and allows for balanced horizontal pleiotropy (Rees et al., 2019). In other words, the IVW multiplicative random-effects model allows each SNP to have different mean effects due to horizontal pleiotropy if the average effect of the pleiotropy is zero (so-called balanced horizontal pleiotropy) (Hemani et al., 2018b). The fixed-effect model does not account for heterogeneity because it assumes that all the SNPs provide exact causal estimates, meaning none exhibit horizontal pleiotropy (Burgess et al., 2019). In addition, the IVW multiplicative random-effects model was also preferred to the additive random-effect model because the latter upweights outlying estimates, which are likelier to represent pleiotropic variants (Burgess et al., 2019).

IVW has the greatest statistical power to detect causal effects, but if an SNP exhibits horizontal pleiotropy, the IVW effect estimate is liable to be biased. Therefore, I undertook several sensitivity analyses to assess the validity of the IVW results. I mainly focused on sensitivity analyses that relax the assumption of no horizontal pleiotropy and have different assumptions to the IVW and to each other (Hemani et al., 2018b). In contrast to the IVW method, these methods have a lower power to detect a causal effect. Thus, these sensitivity analyses were mainly to assess whether the magnitude and direction of effect estimates were consistent across methods and cannot replace the IVW (as the main estimator) unless there is evidence of horizontal pleiotropy.

2.3.2.3. Sensitivity analyses

2.3.2.3.1. MR–Egger method

MR-Egger is an extension of IVW that also combines Wald ratio estimates from each SNP in the random-effects model and uses the inverse SE of the SNP-outcome association estimates as weightings to obtain the casual estimate of the exposure on the outcome. Unlike IVW, MR–Egger regression yields an unbiased estimate, even if all the genetic variants exhibit horizontal pleiotropy (Burgess and Thompson, 2017). This would be the case if the horizontal pleiotropy was independent of the SNP's effects on the exposure, referred to as the InSIDE (instrument strength independent of direct effect) assumptions (Burgess and Thompson, 2017).

The MR-Egger estimates consist of two parts: the causal estimate of the exposure on the outcome adjusted for horizontal pleiotropy and the intercept term, which is interpreted as an estimate of the average pleiotropic effect across the genetic variants (i.e. the average direct effect of an SNP with the outcome) (Burgess and Thompson, 2017, Bowden et al., 2015). When the horizontal pleiotropy effect is absent, the intercept term will be equal to zero, and the estimates from MR-Egger and IVW will be consistent (Bowden et al., 2015). However, in the existence of horizontal pleiotropy effects, the intercept term will differ from zero, and the effect estimates from MR-Egger and IVW will be different (Bowden et al., 2015). More details about the MR-Egger intercept can be found in Section 2.3.2.3.3.2.

The MR-Egger is mainly used as a sensitivity tool and does not replace IVW unless there is evidence of horizontal pleiotropy. This is because, in contrast to IVW, MR-Egger has the lowest power to detect the causal effect, as the SE from MR-Egger is larger than the SE from IVW to account for pleiotropy. Moreover, precision in the MR-Egger method requires SNPs to have a wide range of associations with the exposure (Burgess and Thompson, 2017, Bowden et al., 2016). Thus, when all the genetic variants have the same magnitude of association with the exposure, the MR-Egger p-value would not be significant (Bowden et al., 2016). Furthermore, whilst IVW is less severe to outliers, the MR-Egger is very sensitive to outliers. This means that if one genetic variant has a greater association with the exposure than others, this variant will have a large impact on the effect estimate in the MR-Egger regression and can reverse the sign (i.e. direction) of the MR-Egger effect estimate (Burgess and Thompson, 2017).

2.3.2.3.2. Weighted median

The weighted median is the median of the weighted Wald ratio estimates using the inverse SE of the SNP-outcome association estimates as weights (Bowden et al., 2016). Unlike MR-Egger, the weighted median does not require the InSIDE assumption; it assumes that up to 50% of the weight comes from valid genetic variants and allows the other 50% of the weight of the estimate to come from invalid genetic instruments (Bowden et al., 2016, Burgess and Thompson, 2017). Therefore, the weighted median preserves greater precision in the estimates than MR-Egger (i.e. the weighted median has an SE smaller than the MR-Egger SE) (Bowden et al., 2016). However, the estimate will be biased when a single genetic variant is pleiotropic and contributes more than 50% of the weight to the overall effect estimate or if a group of variants are pleiotropic and, together, they contribute more than 50% of the weight (Au Yeung et al., 2018).

2.3.2.3.3. Pleiotropic assessment

MR studies are vulnerable to pleiotropic effects because, in the human genome, a single genetic variant commonly influences multiple traits (Hemani et al., 2018a). Thus, assessing whether genetic variants included in the analysis have pleiotropic effects on the outcome is a crucial step in estimating the causal effect reliably. To do so, I used three statistical tests that are widely employed in MR analyses: Cochran's Q statistic, the I^2 statistic and the MR-Egger intercept.

2.3.2.3.3.1. Cochran's Q and I^2 tests

One of the major sources of heterogeneity in MR is pleiotropy; thus, Cochran's Q statistic represents a diagnostic tool that is used to test whether differences between Wald ratio estimates from each SNP are the result of chance variations or represent violations of no horizontal pleiotropy assumptions. Cochran's Q statistic is a chi-squared distribution with degrees of freedom equal to the number of SNPs minus one under the null hypothesis that all the genetic variants have the same effect (i.e. homogeneity) (Greco et al., 2015). If the Cochran's Q statistic value is higher than the degree of freedom and the p-value of Cochran's Q < 0.05, the null hypothesis of homogeneity would be rejected, indicating potential heterogeneity in the effect estimate (Greco et al., 2015). Cochran's Q statistic has the greatest statistical power when the number of SNPs included in the analysis is large and the lowest when the number of SNPs is small (Greco et al., 2015). To overcome this problem, I used the I^2 statistic proposed by Higgins and colleagues to estimate the magnitude of the heterogeneity (Greco et al., 2015).

The I^2 statistic or index describes the percentage of the total variation in the Wald ratio estimates explained by heterogeneity rather than sampling errors, independent of the number of SNPs (Greco et al., 2015). To facilitate the interpretation of I^2 statistics, Higgins et al. suggest a method for interpreting the degree of heterogeneity, according to which I^2 values $< 25\%$ suggest slight heterogeneity, values >

50% indicate significant heterogeneity and values between 25% and 50% suggest moderate heterogeneity (Higgins et al., 2003).

2.3.2.3.3.2. MR-Egger intercept

I also used the MR-Egger intercept term, which is a straightforward approach to detecting pleiotropy in MR analysis. Under the InSIDE assumptions, if the intercept term is different from zero and the p-value of the intercept < 0.05, this indicates a horizontal pleiotropy (Bowden et al., 2015). Meanwhile, if the intercept is zero and the p-value of the intercept > 0.05, then the pleiotropy effect across the genetic variants is balanced (Bowden et al., 2015). Once the pleiotropy is balanced, heterogeneity among the causal Wald ratio estimates of a set of SNPs could arise due to other factors which assumed not to cause bias in the MR estimate. For example, in the case of binary data, in which the SNP-outcome associations were measured on the odds ratio scale, the non-collapsibility of the odds ratio will introduce heterogeneity into the estimate (Hemani et al., 2018a). Heterogeneity may also arise due to the use of genetic variants that influence the exposure of interest via different biological mechanisms and hence are likely to influence the outcome for which the exposure is a cause at the different magnitude, leading to differences between the Wald estimates, and therefore heterogeneity (Burgess et al., 2019, Foley et al., 2021). Meanwhile, variants that influence the exposure in a similar biological mechanism are likely to have similar Wald estimates, thus no heterogeneity (Foley et al., 2021).

2.3.2.3.3.3. IVW MR-Radial

When there is evidence of horizontal pleiotropy in the estimate, I performed IVW MR-radial to identify the pleiotropic variants that have the largest contribution to heterogeneity (Bowden et al., 2018b). To estimate the contribution of each SNP to the heterogeneity accurately, I used the so-called 'modified second-order weighting' (Bowden et al., 2018b). This weight has been recommended, as it does not inflate the type I error rate of Cochran's Q statistic and maintains the statistical power for

detecting heterogeneity (Bowden et al., 2018b). If an individual SNP's Q contribution is extreme (Cochran's Q p-value < 0.05), MR-radial flags this SNP as an outlier (Bowden et al., 2018b). These outlier SNPs will be removed and then the exposure-outcome causal association will be re-estimated. Although the outlier removal strategy can certainly reduce heterogeneity and bias in MR estimates, caution must be taken because this strategy would reduce the SE and may lead to over-fitting (Hemani et al., 2018a).

2.3.2.4. Further Analysis

2.3.2.4.1. Multivariable MR analysis

Indeed, in some cases, finding variants solely associated with the exposure of interest may be difficult, and an MR analysis cannot be performed without considering pleiotropic variants (Burgess and Thompson, 2015). One method that can be used in this case is multivariable MR analysis (MVMR) (Burgess and Thompson, 2015). MVMR is an extension of univariable MR proposed to analyse genetic variants associated with multiple exposures in one model to account for pleiotropy. Univariable MR and MVMR estimate different causal effects of the exposure of interest on the outcome. In general, MVMR estimates the independent direct effect—i.e. the direct effect comprising all the other pathways from the exposure of interest to the outcome that does not act through the other exposures included in the model (Burgess and Thompson, 2015, Sanderson, 2021). Meanwhile, univariable MR estimates the total effect—i.e. the total effect is the effect of the exposure of interest on the outcome that operates through the other risk factors (mediators) included in the model (Burgess and Thompson, 2015, Sanderson, 2021).

Since the lipid traits and the anthropometric-related traits share common genetic variants, resulting in overlap of the genetic instruments, I used the MVMR through IVW (Burgess and Thompson, 2015) to account for the potential pleiotropic effect and to determine whether several of the exposures affect MS through the same pathway or whether these traits have independent effects.

For the purposes of clarification, I used BMI, FP, FM and MS risk as an example to show how the genetic instruments have been selected for MVMR analysis through the following steps (**Figure 2.3**):

1. If there was no evidence of heterogeneity or pleiotropy in the estimates from univariable analyses, the same set of genetic instruments (266 SNPs) of the risk factor of interest (BMI) was used to extract instruments for other exposure (FP and FM) that I wish to include in the MVMR model.
2. If there was evidence of residual heterogeneity or unbalanced pleiotropy in the estimates from univariable analyses, the MR-radial method was first used to identify and remove pleiotropic variants (13 SNPs) to ensure the validity of MR results. The remaining genetic instruments (253 SNPs) were then used to construct the final list of genetic variants that I wish to extract from the other exposures included in the model.
3. As a requirement in MVMR analysis, each genetic instrument should be strongly (p-value < 5×10^{-8}) associated with the risk factor of interest (BMI) but does not require each genetic instrument to be strongly associated with the other exposure included in the model (FP and FM) (Burgess and Thompson, 2015).
4. The *mv_multiple()* function, which is available in the Two-SampleMR R package, was then used to obtain the direct effect of the risk factor of interest on the outcome (Hemani et al., 2018b). In general, this function regressed the genetic instruments for each exposure against the outcome (MS risk) together, weighting for the inverse variance of the outcome (Hemani et al., 2018b).

In the interpretation of the MVMR-IVW results, the estimate with a p-value ≤ 0.05 indicates that the exposure of interest affects MS through a pathway independent of other exposures considered in the model (Sanderson, 2021).

The attenuation of the direct effect (MVMR-IVW p-value > 0.05) compared to the corresponding univariable MR results (IVW p-value < 0.05) indicates that the effect of the exposure of interest on MS is mediated through the other exposure considered in the model (Sanderson, 2021).

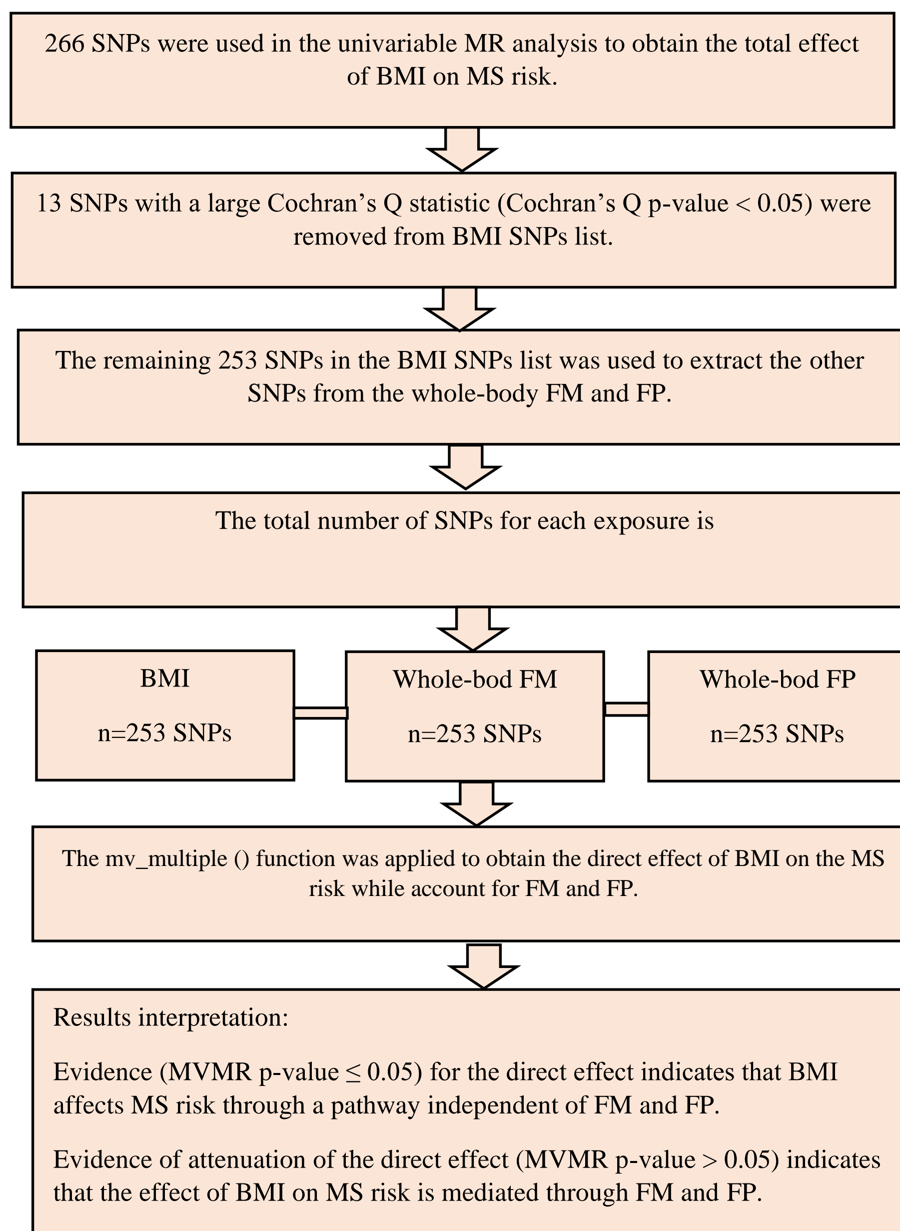


Figure 2.4: A detailed flow chart showing the steps involved in genetic instrument selection and the interpretation of the results for MVMR analysis using BMI, FM, FP and MS risk as an example. Abbreviations: BMI, body mass index; FM, fat mass; FP, fat percentage; MS, multiple sclerosis; MVMR, multivariable Mendelian randomization.

2.3.2.4.2. Bidirectional MR

The causal association between two traits may be bidirectional, meaning that the exposure influences the outcome or that the outcome influences the exposure. For example, MR studies reported that smoking reduces BMI and higher BMI increases smoking (Carreras-Torres et al., 2018, Winsløw et al., 2015). Therefore, I carried out MR analyses in the opposite direction (i.e. switching around the exposure and outcome) to study whether the liability to MS produces a predisposition to any of the risk factors (lipids, anthropometric-related measures, stroke) that have been used in this thesis to assess their causal role on MS risk (so-called bidirectional MR). With MS risk as exposure and the risk factors of interest as the outcome, I followed the same steps for genetic variant selection and MR analysis pointed out above to obtain the causal estimates for this bidirectional MR analysis.

2.3.2.5. Causal effect estimation using correlated variants

The eQTL data are among the datasets used in this thesis. Since SNPs in eQTL data are highly correlated, using a stringent clumping threshold ($r^2 < 0.001/0.01$) would result, in most cases, in only one SNP, where it is not possible to use IVW or MR-Egger methods that require at least three SNPs. Therefore, a liberal (e.g. $r^2 < 0.4$) clumping threshold was used with gene expression data to allow for more SNPs to be included in the model, which would likely lead to a more reliable analysis (Burgess et al., 2016). Further, the inclusion of multiple SNPs in partial LD can explain a greater proportion of variance in the exposure, thus leading to a more powerful MR (Burgess et al., 2016). I then carried out MR analysis using IVW and MR-Egger, which were extended to account for the correlations between SNPs via incorporating the LD matrix (Burgess et al., 2016). Estimates of SNP correlations have been obtained from the European 1,000 Genomes Project via the *dat_to_MRInput* function.

2.3.2.6. Multiple test correction

The adjustment for multiple testing has become standard in areas such as genomics, neuroimaging, proteomics, psychology, and economics (Korthauer et al., 2019) . Thus, the false discovery rate (FDR) (Benjamini and Hochberg, 1995) was used in this thesis to adjust the p values computed for the IVW method. This multiple test correction has been conducted separately for the analyses of each chapter, in which exposures with significant adjusted p values ≤ 0.05 were defined as exposures with potential evidence of a causal effect.

The concept of the FDR was proposed by Benjamini and Hochberg to control for the proportion of false-positive conclusions when multiple hypotheses are simultaneously tested in order to restrict the total number of false discoveries (Groenwold et al., 2021, Korthauer et al., 2019, Benjamini and Hochberg, 1995). The FDR has emerged as a popular and powerful approach, as it is the least conservative method that provides a good balance between discovering statistically significant effects and the limitations of false positives (Korthauer et al., 2019, Menyhart et al., 2021). Other techniques, such as Bonferroni correction, are highly conservative, and, while they still control for the probability of any false positives, the power to detect true positives is greatly reduced (Korthauer et al., 2019). In practice, the Benjamini-Hochberg critical value is calculated by first ranking the p values in ascending order and then multiplying each p value by the total number of tests and dividing by the p value's individual rank.

2.3.2.7. Reporting the findings

In the absence of evidence of horizontal pleiotropy (MR-Egger intercept p value > 0.05), I used the IVW estimates as the most reliable indicator to report the MR results of the underlying causal relationship. This is because IVW assumes that genetic variants are valid instruments (balanced horizontal pleiotropy). If there was evidence of horizontal pleiotropy, I used MR-Egger to report the MR results as designed to account for horizontal pleiotropic effects (Burgess and Thompson, 2017).

2.3.2.8. Specialist software for MR statistical analyses

I used R software with the aid of R packages to conduct all the above MR methods (Team, 2017). R is a free software environment that allows users to conduct statistical and graphics projects. R packages that have been used to conduct MR analysis include the 'TwoSampleMR' (Hemani et al., 2018b) , 'RadialM' (Bowden et al., 2018b), 'Biomart' (Durinck et al., 2005) and 'corrplot' (Wei et al., 2017).

Chapter 3 *Exploring the Role of Lipids and Statins on Multiple Sclerosis Risk and Severity*

Statement of contribution

I developed the idea, wrote and executed the analysis scripts used in this chapter myself. Nicholas Wood suggested the idea of this chapter. Chris Finan, Catherine Storm, Amand Schmidt, Demis Kia, Sandesh Chopade, Rachel Coneys, Nicholas Wood abd Aroon Hingorani, contributed to the interpretation of my results.

3.1. Introduction

A significant amount of attention has been drawn to statins not only because they lower blood cholesterol levels and protect against cardiovascular morbidity but also due to their pleiotropic effects beyond their lipid-lowering properties. Indeed, recent evidence derived from clinical and experimental animal models of autoimmune diseases has shown that statins exert immunomodulatory and anti-inflammatory effects that may be beneficial in autoimmune diseases such as MS (Greenwood et al., 2006, Weber et al., 2007). Many of these pleiotropic effects are predominantly ascribed to their capacity to inhibit the isoprenylation (also known as prenylation or lipidation) of Rho small guanosine triphosphatases (GTPases; also known as small G-proteins) (Takemoto and Liao, 2001, Cai et al., 2015, Wang et al., 2008, Neuhaus et al., 2004).

Recent findings from the phase 2 MS-STAT trial (a randomised, placebo-controlled trial) showed that a high dose of simvastatin (80 mg per day) led to a significant drop in brain atrophy (by 43%) and disability progression among 140 patients with SPMS over two years (Chataway et al., 2014). Observational and molecular studies suggest a role for lipid metabolism in MS pathogenesis as a biomarker for predicting MS disease activity and progression (Pineda-Torra et al., 2021, Chandra and Xu, 2016). It has been shown that there is an association between dyslipidaemia, namely elevated levels of circulating LDL-C and TG and reduced levels of HDL-C, and greater MS disease activity (i.e. worsening disability and new magnetic resonance imaging lesions in MS) and this may also contribute to the pathogenesis of MS (Weinstock-Guttman et al., 2011, Tettey et al., 2014),(Zhornitsky et al., 2016). These findings raise important questions about the nature of the relationship between lipids and MS and whether statins can be used as part of a prevention strategy to reduce the risk and/or severity of MS.

A previous MR analysis used SNPs within the 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR) gene region to mimic the effects of statins on the risk of MS developing via *HMGCR* inhibition. This study revealed no causal link between these SNPs and MS risk, suggesting that statins have no effect on MS risk (Yang and Schooling, 2021). *HMGCR* is the target for statins; therefore, it is not surprising that MR studies focus on *HMGCR* to mimic the effects of statins. Nevertheless, by only targeting *HMGCR*, these studies examined the cholesterol-lowering effect only and may have missed observing the statins' pleiotropic effects. Furthermore, the effect of statins on MS severity has not yet been established. To address this knowledge gap, I adopted an MR approach to genetically mimic both cholesterol-dependent and cholesterol-independent effects of statins to explore whether statins' effects on MS , if any, are mediated by lowering cholesterol or are independent of cholesterol. In particular, a set of genes involved in cholesterol biosynthesis and Rho GTPases were included in the MR analyses for this purpose. To my knowledge, this is the first time that the MR approach has been applied to genetically mimic the effects of stains through multiple pathways. In addition, the causal role of lipids in both MS risk and severity is addressed in this chapter.

3.1.1. Limitation of genetically mimicking the effects of statins via HMGCR

MR has been widely employed to explore the repurposing opportunities and the side effects of statins on disease outcomes. Typically, SNPs in *HMGCR* gene are used that affect LDL-C levels to mimic the statins' biological effects. Interestingly, MR findings generated using such SNPs have been broadly consistent in terms of the therapeutic performance of statins in clinical trials. For example, MR results showed that *HMGCR* SNPs had a causal effect on reducing the risk of CHD (Ference et al., 2015). This finding was in line with those of several RCTs that revealed the clinical benefits of statin therapy in reducing LDL-C levels and therefore reducing mortality and morbidity in adults with CHD (Wilt et al., 2004). However, the observed effect of statins in clinical trials on some conditions, such as neurocognitive and autoimmune-related conditions, were not replicated in MR studies when *HMGCR* variants were used (Liu et al., 2021a).

For example, MR studies found no evidence of a causal role for *HMGCR* variants and genetic predisposition to increased blood lipid levels (LDL-C, HDL-C, TG and total cholesterol) on the risk of Alzheimer's disease (AD) (Proitsi et al., 2014, Østergaard et al., 2015, Benn et al., 2017, Williams et al., 2020). In contrast, numerous epidemiological and small clinical studies have suggested that statins are associated with a reduced risk of AD (Sparks et al., 2008, Wolozin et al., 2000, Poly et al., 2020). This raises the question of what is causing the discordance in the evidence generated in MR studies, RCTs and epidemiological studies regarding the role of statins on AD.

Indeed, as mentioned above, MR studies have focused on genetically mimicking the effects of statins via the use of *HMGCR* variants, because *HMGCR* is the main target for statins, and have ignored the other genes downstream of the *HMGCR* pathway that contribute to statins' pleiotropic effects. Therefore, the lack of evidence in MR analyses possibly implies that the statins effect has most likely been missed by only focusing on genetically mimicked statins via *HMGCR* variants.

Evidence generated in MR studies has shown that LDL-C is not causally related to AD (Proitsi et al., 2014, Østergaard et al., 2015). This finding plausibly explains the absent of causal link between *HMGCR* variants and AD risk because it implies that if LDL-C is not a risk factor, then using a drug intended to reduce the cholesterol levels would not be an effective therapeutic strategy to reduce AD risk. This finding also suggests that statins may act on AD via other pathways independent of the lowering of cholesterol. This hypothesis has been confirmed in experimental studies using human immune cells and mouse microglial cells to understand the beneficial effect of statins in reducing the risk of AD (Cordle and Landreth, 2005, Cordle et al., 2005). The findings revealed that statins exert anti-inflammatory effects by preventing isoprenylation of members of the Rho GTPase family (RhoA and Rac1), resulting in inhibition of the beta amyloid-mediated inflammatory response (Cordle and Landreth, 2005, Cordle et al., 2005).

HMGCR variants seem ideal targets for MR analysis when LDL-C is an overt risk factor. Such in case of CHD in which a high circulating LDL-C level is an established risk factor for CHD; therefore, the ability of statins to reduce LDL-C is an effective mechanism for CHD prevention (Ference et al., 2015). Thus, MR analyses provided evidence of the causal association between *HMGCR* SNPs and reduction in CHD risk (Ference et al., 2015).

This evidence motivated me to study the beneficial effects of statins on MS and identify the mechanism by which statins mediate their effect on MS. In the following section, I provide an overview of the cholesterol- dependent and cholesterol-independent pathways and the pleiotropic effects that statins provide via Rho GTPases.

3.1.2. Statin activity via the cholesterol-dependent pathway

Statins competitively inhibit *HMGCR*, the rate-limiting enzyme of the cholesterol biosynthesis pathway, which is also known as the mevalonate (MVA) pathway (Liao, 2002, Cai et al., 2015). Blockade of *HMGCR* via statins is accompanied by an increase in hepatic LDL receptor that promotes uptake and clearance of cholesterol from the bloodstream and ultimately reduces plasma LDL-C and TG levels (Liao, 2002, Cai et al., 2015). Also, using statins to block *HMGCR* inhibits the synthesis of isoprenoids, such as farnesyl pyrophosphate (FPP) and geranylgeranyl pyrophosphate (GGPP), which is an essential lipid attachment for anchoring Rho GTPases to the membrane (**Figure 3.1**) (Liao, 2002). Therefore, members of the Rho GTPase family that undergo isoprenylation are important targets for mediating the biological effects of statins (Liao, 2002).

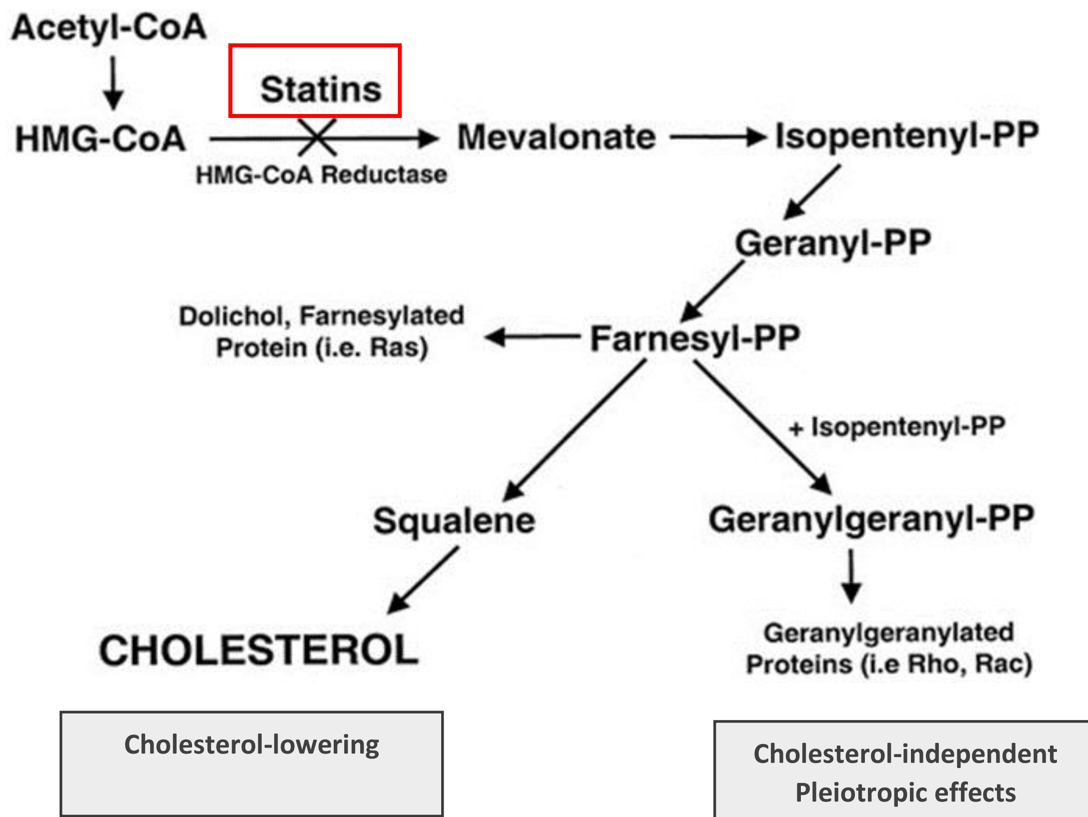


Figure 3.1: The cholesterol biosynthesis pathway. 3-Hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR) inhibition by statins reduces the synthesis of cholesterol and isoprenoids. Adapted from (Cai et al., 2015).

3.1.3. Statin activity via the cholesterol-independent pathway (pleiotropic effects via Rho GTPases)

The Rho family of GTPases is a subgroup of the Ras superfamily of small GTP-binding proteins (Azzarelli et al., 2015). Although the most extensively studied members of the Rho family are Cdc42, Rac1 and RhoA, this family includes 20 members that are divided into two classes: classical and atypical. The classical Rho GTPases are further subdivided into four subgroups (Rho, Rac, Cdc42 and Rif), and the atypical Rho GTPases are subdivided into four subgroups (Rnd, RhoUV, RhoH and RhoBTB) (Azzarelli et al., 2015, Bayo et al., 2021) (**Table 3.1**).

The classical Rho GTPases act as molecular switches, cycling between the active guanosine triphosphate (GTP)-bound and inactive guanosine diphosphate (GDP)-bound states to regulate physiological cell processes, such as cell division, migration, transcription, cell-cycle progression and apoptosis (Bayo et al., 2021). The activity of classical Rho proteins can be tightly regulated by lipid modification (e.g. with isoprenoids), guanine nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs) and guanine nucleotide dissociation inhibitors (GDIs) (Bayo et al., 2021, Hodge and Ridley, 2016) (**Figure 3.2**).

For classical Rho proteins to be functional and to localise to internal membranes within eukaryotic cells, these proteins must be isoprenylated. During isoprenylation, a post-translational modification process, isoprenoids (GGPP and FPP) produced from the cholesterol biosynthesis pathway are attached to the C-terminal cysteine residues of Rho proteins (Greenwood et al., 2006, Ridley, 2006). After isoprenylation, the Rho proteins localise to a target cell membrane and are activated by GEFs that facilitate the exchange of GDP for GTP. This enables them to pass on signals to corresponding downstream molecules and mediate numerous cellular functions in different cell types (Hodge and Ridley, 2016, Bayo et al., 2021). Finally, the Rho proteins interact with GAPs that hydrolyse GTP to GDP, thereby inactivating the Rho proteins (Hodge and Ridley, 2016, Bayo et al., 2021). When the Rho proteins are inactivated (GDP-bound form), GDIs extract them from the membrane and sequester the proteins in the GDP-bound form into the cytosol, preventing them from anchoring to membranes or being activated by GEFs (Hodge and Ridley, 2016, Bayo et al., 2021). It should be noted that GDIs can only bind to isoprenylated Rho proteins (Hodge and Ridley, 2016) (**Figure 3.2**).

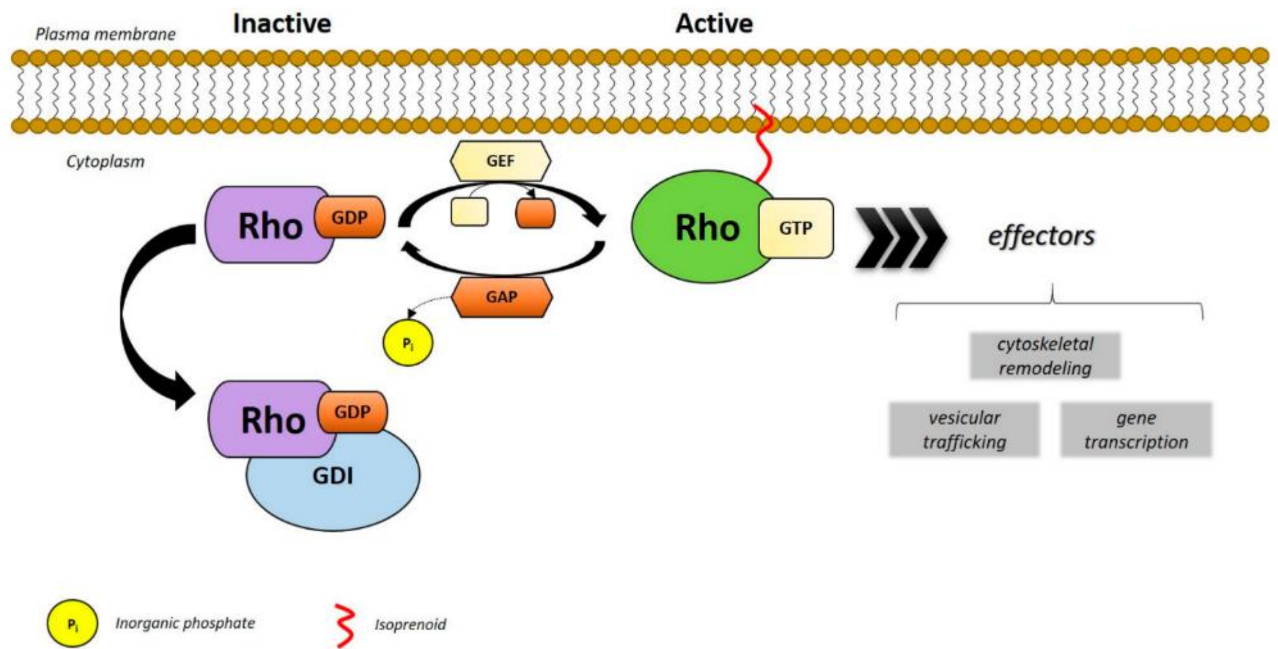


Figure 3.2: The regulation of Rho proteins. Rho guanosine triphosphatases (GTPases) cycle between the inactive guanosine diphosphate (GDP)-bound and active guanosine triphosphate (GTP)-bound forms, and this is regulated by guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs) (Pradhan et al., 2021). Rho GTPases bind to effector molecules to generate a downstream response in their active form. Post-translational modifications, such as the attachment of isoprenoid groups (e.g. geranylgeranyl pyrophosphate; GGPP), allow the proteins to localise to the plasma membrane (Pradhan et al., 2021). Reproduced from (Pradhan et al., 2021).

Atypical Rho subfamily proteins, including Rnd, RhoUV, RhoH and RhoBTB (**Table 3.1**), do not follow the classical scheme of Rho activation (GDP/GTP cycling) and therefore do not require GEFs and GAPs (Azzarelli et al., 2015, Hodge and Ridley, 2016). They are constitutively GTP-bound, they do not hydrolyse GTP to GDP and their activities can be regulated by transcriptional regulation or phosphorylation rather than via the classical cycling process (Azzarelli et al., 2015, Hodge and Ridley, 2016). This is either because the GTPase domain of the atypical proteins harbours high intrinsic nucleotide exchange activity or because they have a different set of amino acids that prevents GTPase activity (Azzarelli et al., 2015, Hodge and Ridley, 2016).

Table 3.1: Members of the Rho GTPase family

Rho type	Rho subfamilies	Rho members	Identity	Localization	Post-translational modifications
Classical Rho GTPases	Rho	RhoA		PM and cytosol	GGPP and PH
		RhoB	84%	PM and endosomes	GGPP, FPP and P
		RhoC	92%	PM and cytosol	GGPP
	Rac	Rac1		PM	GG PP
		Rac2	92%	PM and cytosol	GGPP
		Rac3	93%	PM and endomembranes	GGPP
		RhoG	72%	PM and endosomes	GGPP
	Cdc42	Cdc42		PM and Golgi	GGPP
		TC10 (RhoQ)	62%	PM and perinuclear	FPP and P
		TCL (RhoJ)	55%	PM and endosomes	FPP and P
	Rif	Rif (RhoF)		PM	GGPP
		RhoD	50%	PM and endosomes	GGPP
Atypical Rho GTPases	Rnd	Rnd1		PM	FPP
		Rnd2	51%	endosomes and cytosol	FPP
		Rnd3 (RhoE)	58%	PM, Golgi and cytosol	FPP and PH
	RhoUV	RhoU (Wrch, Chp2)		PM and endomembranes	P
		RhoV (Chp, Wrch2)	54%	PM and endomembranes	P
	RhoH	RhoH (TTF)		None known	GGPP
	RhoBTB	RhoBTB1		Vesiculard	None known
		RhoBTB2	66%	Vesiculard	None known

The first five columns were reproduced from (Azzarelli et al., 2015), and the last two columns were reproduced from (Ridley, 2006). The column identity refers to the percentage identity in the amino acid sequence of a specific Rho GTPase compared with the first member of the corresponding subfamily (Azzarelli et al., 2015). Abbreviations: GGPP, geranylgeranyl pyrophosphate; FPP, farnesyl pyrophosphate; P, palmitoylation; PH, phosphorylation; PM, plasma membrane.

3.1.4. Effects of statins mediated by Rho GTPases

Statins exert effects via Rho GTPases by two distinct mechanisms: preventing Rho proteins from localising to the membrane localisation and loading Rho proteins with GTP. By inhibiting *HMGCR*, statins prevent the synthesis of isoprenoid intermediates and the subsequent isoprenylation of Rho GTPases (Rikitake and Liao, 2005). This leads to the inhibition of Rho protein translocation to the plasma membrane and thus prevents the activation of their downstream effectors (Rikitake and Liao, 2005). For example, studies with a mouse model of ischaemic stroke and human endothelial cells showed that statins increase endothelial nitric oxide synthase (eNOS) expression, an important mediator of vascular homeostasis and blood flow, by inhibiting RhoA membrane translocation and activity and that the effect is independent of serum cholesterol levels (Endres et al., 1998, Laufs et al., 2000, Laufs and Liao, 1998).

The second mechanism by which statins exert effects via Rho GTPases is GTP loading, which is the conversion of Rho proteins to their active form (GTP-bound). Inhibition of isoprenoid biosynthesis by statins also results in disruption of GDI–Rho GTPase binding, which provides a potential mechanism for GTP loading of the cytosolic Rho proteins (Zhu et al., 2013, Cordle et al., 2005). As mentioned in section 3.13, GDIs only bind to isoprenylated Rho proteins; thus, in the absence of these isoprenoid intermediates, GDIs cannot bind to Rho proteins, allowing them to be constitutively active (GTP-bound). In support of this notion, Zhu et al. demonstrated that applying simvastatin to pre-treatment HCT116 colorectal cancer cells increased GTP loading of cytosolic Rac1 and RhoA and decreased their membrane translocation (Zhu et al., 2013). The authors attempted to understand the non-canonical activation of GTP-bound Rac1 and RhoA in the wake of simvastatin treatment by examining the interaction of Rho proteins with a GDI (Zhu et al., 2013). They demonstrated that because simvastatin treatment blocked isoprenoid synthesis, the cytosolic GDI could not interact with Rac1 and RhoA, suggesting a mechanism for enhanced GTP loading of cytosolic Rho GTPases and proposing that it is functionally relevant to apoptosis (Zhu et al., 2013).

3.2. Aims and hypothesis

Using the MR approach, the work in this chapter was designed to assess the following questions:

1. The causal role of genetic predisposition to increased major plasma lipid fractions (HDL-C and TG) in MS risk and severity.
2. Whether genetic predisposition to increased MS risk influences HDL-C, LDL-C or TG levels (i.e. reverse causation).
3. The role of statins in the prevention of MS development and in delaying the progress of MS severity via genetically mimicking the biological effects of statins via the cholesterol-dependent and cholesterol-independent pathways.
    I. The cholesterol-dependent pathway was studied by (a) examining the causal role of the change in the blood expression levels of 25 genes (including the *HMGCR* gene, which is the statins' target) that encode proteins involved in cholesterol biosynthesis, and (b) examining the causal role of genetically predicted LDL-cholesterol, given that LDL-C is a relevant prognostic factor for assessing the degree of *HMGCR* inhibition (Carter et al., 2020).
    II. The cholesterol-independent pathway was studied by examining the causal role of the change in the blood expression levels of 20 genes that encode Rho GTPase family members.

I tested 2 hypotheses to examine whether statins influence MS through cholesterol-dependent or cholesterol-independent pathways:

1. If statins causally affect MS via lowering blood cholesterol levels, then I expect to see:
    a. A statistically significant causal estimate for MR analyses involving LDL-associated SNPs.
    b. A statistically significant causal estimate for MR analyses involving SNPs of *HMGCR* and any other downstream genes involved in cholesterol biosynthesis.
2. If statins causally affect MS and the effect is not dependent on cholesterol, I expect to see a statistically significant causal estimates for MR analyses involving SNPs of Rho GTPases.

3.3. Method and results

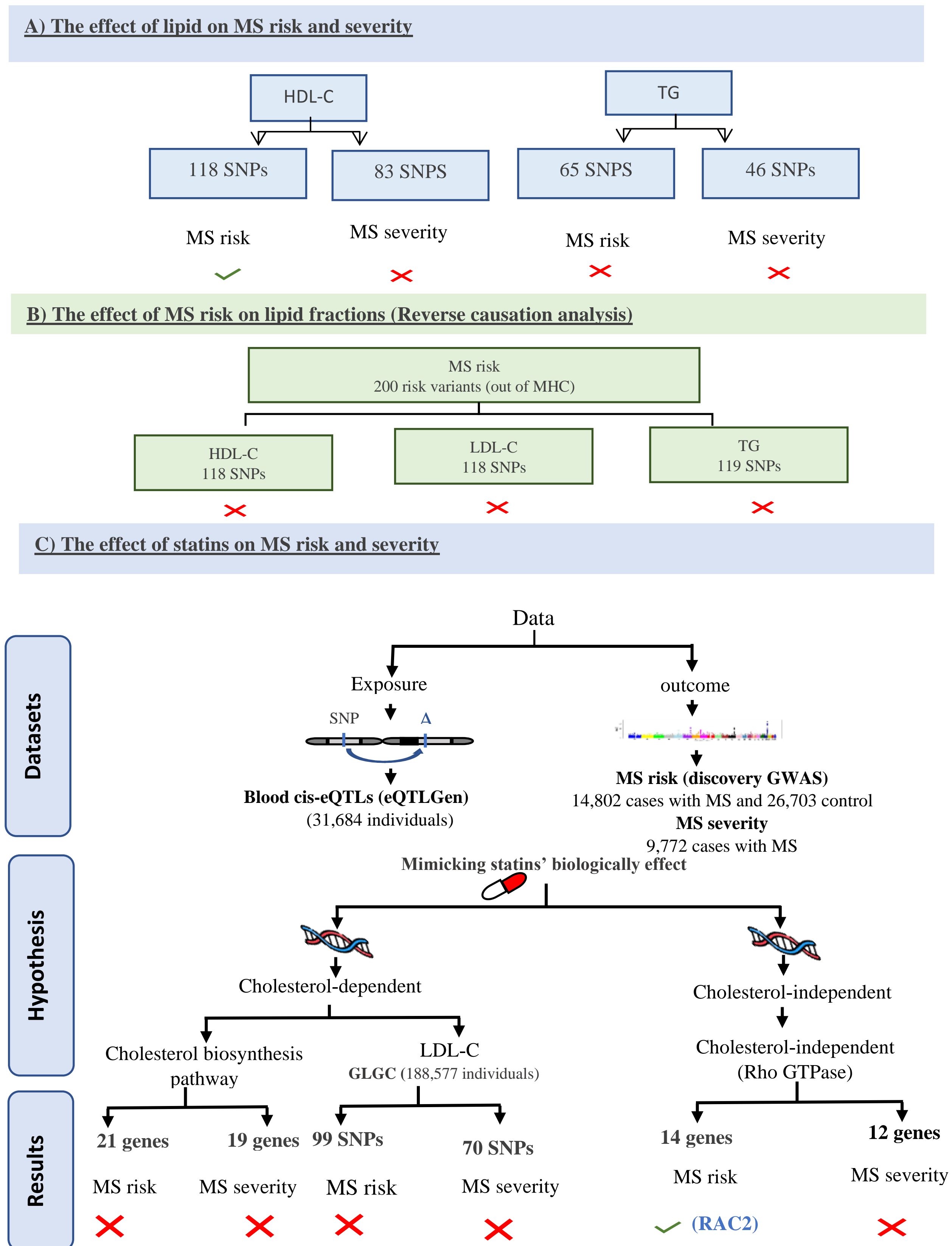


Figure 3.3: A flow diagram summarising this chapter's method and results**.** The cross symbol indicates that there is no causal association, while the tick symbol indicates that there is casual association.

3.3.1. Genetically predicted HDL-C associated with increased MS risk but not severity

To investigate whether blood lipids are causally associated with MS risk and severity, genetic associations with lipid exposures (HDL-C, LDL-C and TG) were obtained from the GLGC (methodology chapter section 2.1.21). All the selected SNPs were associated with target lipid fractions at p-values $< 5 \times 10^{-8}$ and clumped at an LD threshold value of $r^2 < 0.01$. With the lipid-associated SNPs recorded as exposures, I then obtained the corresponding effect estimates for MS risk from the IMSGC (methodology chapter section 2.2.3) and for MS severity from the corresponding author of the original publication (Professor Jacob McCauley; methodology chapter section 2.1.3) as outcomes.

Table 3.2 presents the number of SNPs, the explained variance (R^2) and the mean F-statistics for each lipid trait. Since the F-statistics for the independent SNPs were all more than 10, bias due to weak instruments is negligible in this data.

Table 3.2: Sample characteristics of the lipid traits

Lipid trait	Lipid-MS risk			Lipid-MS severity		
	No. of SNPs	R^2 (%)	Mean F-statistics	No. of SNPs	R^2 (%)	Mean F-statistics
HDL-C	118	9	124	83	6.9	139
LDL-C	99	11.6	159	70	7.7	156
TG	65	6.3	159	46	5.2	189

HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TG, triglyceride; No. of SNPs, the number of independent genome-wide significant single nucleotide polymorphisms; R^2 (%), approximate variance explained by SNPs in the target trait that expressed in percentage.

MR analysis was performed for each of the lipid fractions in turn, and the results of these analyses are displayed in **Table 3.3** and **Figure 3.4**. The associations between genetically predicted lipid fractions and MS risk are reported as odds ratios (ORs) with the 95% CI per one standard deviation (SD) increase of lipid fraction. For HDL-C, assessment through IVW showed evidence that raised HDL-C is associated with an increase in MS risk. The MR-Egger analysis results replicated this finding and the magnitude of the effect. The heterogeneity was significant (Cochran's Q p-value < 0.05). However, since the MR-Egger intercept indicates a balanced horizontal pleiotropy (p-value > 0.05), this heterogeneity is not due to pleiotropic variants. Instead, it is possibly due to a different SNP–HDL-C influence on MS risk mediated via a different biological mechanism. The MVMR analysis results after adjustment for LDL-C and TG remained broadly consistent with the primary findings in the IVW estimator, which further supported the notion of a causality relationship between HDL-C and MS risk. For LDL-C and TG, there was no evidence for a causal relationship with MS risk found in the IVW, MR-Egger and MVMR estimator results (**Table 3.3**). There was evidence of heterogeneity; however, the MR-Egger intercept test did not provide any evidence of horizontal pleiotropy in these results.

Since the HDL-C results were deemed significant (FDR ≤ 0.05) after multiple testing corrections, the results were assessed for replication using independent HDL-C data from UK Biobank (methodology chapter section 2.2.2.3). The replication result aligned with the initial results, further supporting the significant causal association between HDL-C and MS risk (**Table 3.4**). However, I must acknowledge that there was unbalanced pleiotropy in this analysis. Thus, in this case, the MR-Egger estimate, which is robust to pleiotropic instruments, is more reliable than the IVW estimate. To add more confidence to the replication results, I also used MR-radial to identify and remove the pleiotropic variants and then re-analysed using the remaining variants. Following the exclusion of the pleiotropic variants, there was a dramatic reduction in heterogeneity ($I^2 = 0\%$), but the pleiotropy remained significant, implying that the pleiotropic effect is the same across all variants (Bowden et al., 2017). Together, in the discovery and replication analyses, the MR-Egger and IVW results support the existence of a causal role for HDL-C in MS risk.

Table 3.3: MR analysis of the effect of lipid level on MS risk

Lipid trait	Method	No. of SNPs	OR (95 % CI)	p-value	FDR	pleiotropy assessment			
						Q p-value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p-value
HDL-C	IVW	118	1.144 (1.04,1.26)	7.94E-03	2.38E-02				
HDL-C	MR Egger	118	1.23 (1.02,1.48)	3.03E-02		6.38E-06	40.5	-0.004	3.66E-01
HDL-C adjusted for LDL-C & TG	MVMR	118	1.255 (1.06,1.49)	9.13E-03					
LDL-C	IVW	99	0.996 (0.9,1.1)	9.36E-01	9.36E-01				
LDL-C	MR Egger	99	1.002 (0.85,1.18)	9.82E-01		1.30E-09	52.5	-5.00E-04	9.27E-01
LDL-C adjusted for HDL-C & TG	MVMR	99	1.024 (0.94,1.12)	6.07E-01					
TG	IVW	65	0.921 (0.81,1.04)	2.00E-01	3.00E-01				
TG	MR Egger	65	0.859 (0.7,1.05)	1.36E-01		1.50E-04	43.6	0.0046	3.70E-01
TG adjusted for HDL-C & LDL-C	MVMR	65	1.076 (0.88,1.32)	4.72E-01					

Abbreviations: HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TG, triglyceride; No. of SNPs, the number of independent genome-wide significant single nucleotide polymorphisms; IVW, inverse-variance weighted; MVMR, multivariable MR; OR, odds ratio; CI, confidence interval; Q p-value, Cochran's Q statistic; FDR, false discovery rate; I^2 (%) expresses the level of heterogeneity as a percentage.

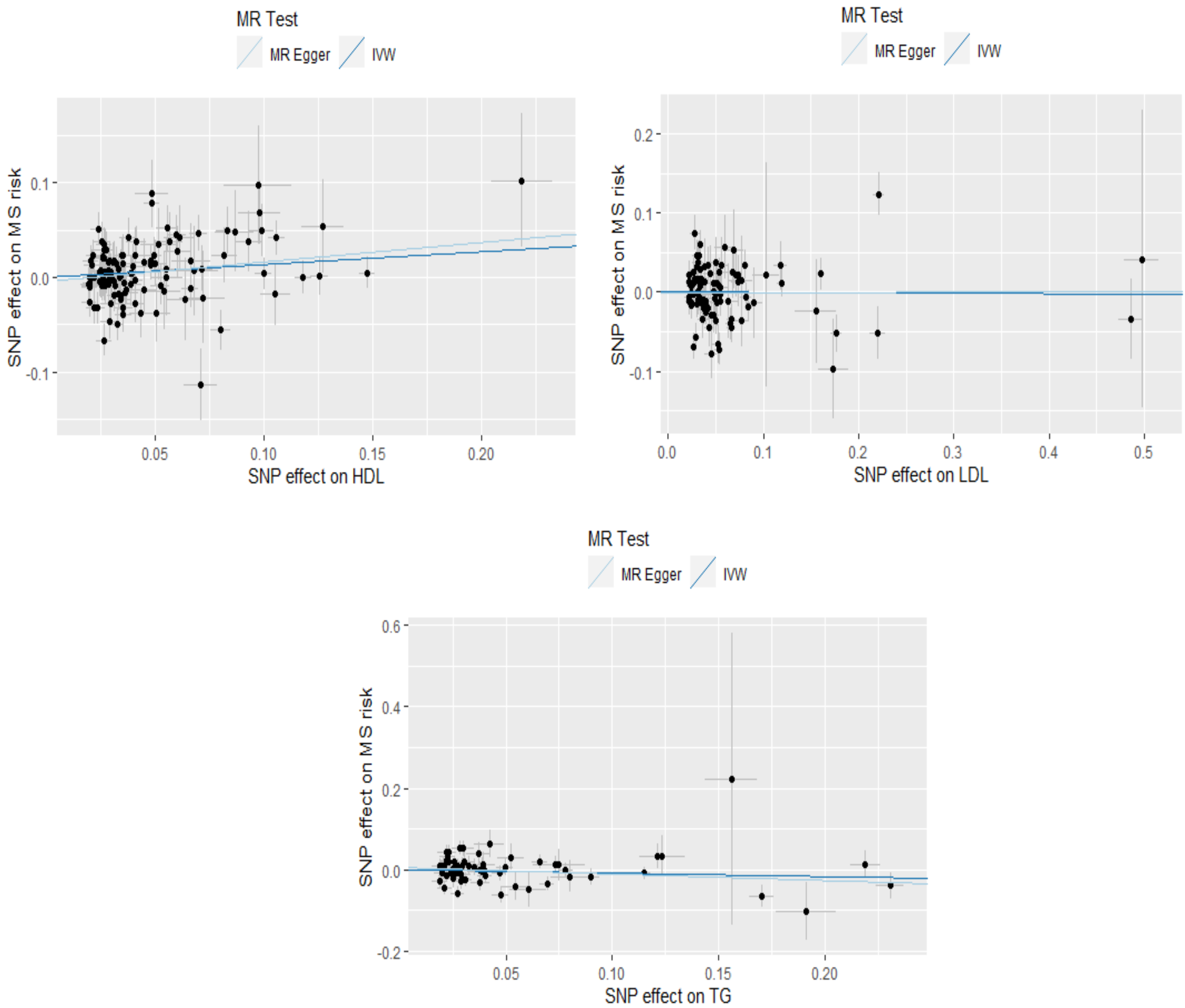


Figure 3.4: Scatter plots for MR analyses showing the causal estimates of the lipid fractions on MS risk. The effect sizes of each genetic variant (with 95% confidence intervals) are represented by black points. The slope of each line shows the estimated MR effect for each method.

Table 3.4: Replication analysis results for the effect of HDL-C on MS risk

	Trait	Method	No. of SNP	OR (95 % CI)	p-value	Q p-value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p-value
Including all the SNPs	HDL-C	IVW	482	1.079(0.99,1.17)	6.63E-02	6.43E-04	18.1		
		MR Egger	482	1.193(1.05,1.36)	7.23E-03	9.17E-04	17.6	-0.0034	4.84E-02
Removing the pleiotropic SNPs		IVW	450	1.098(1.02,1.19)	1.91E-02	1.00E+00	0		
		MR Egger	450	1.209(1.07,1.37)	2.48E-03	1.00E+00	0	-0.0032	4.50E-02

For Abbreviations, see Table 3.3.

The IVW, MR-Egger and MVMR methods were also implemented to assess the lipid influence on MS severity. The results revealed no evidence of HDL-C, LDL-C or TG having a causal role in MS severity. The associations between the genetically predicted lipid fractions and MS severity are given in **Table 3.5**, reported as betas with the 95% CI per 1-SD increase of lipid fractions and graphically represented in **Figure 3.5**. No evidence of heterogeneity or pleiotropy was detected in this analysis**.**

Table 3.5: MR analysis of the effect of lipid level on MS severity

Trait	Method	No. of SNP	Beta (95 % CI)	p-value	FDR	Pleiotropy assessment			
						Q p-value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p-value
HDL-C	IVW	83	-0.155 (-0.33,0.02)	8.47E-02	2.54E-01				
HDL-C	MR Egger	83	-0.039 (-0.49,0.41)	8.67E-01		1.00E+00	0	-0.0067	5.46E-01
HDL-C adjusted for LDL &TG	MVMR	83	-0.059 (-0.4,0.28)	7.35E-01					
LDL-C	IVW	70	-0.091 (-0.3,0.12)	4.02E-01	6.03E-01				
LDL-C	MR Egger	70	-0.019 (-0.43,0.4)	9.29E-01		8.05E-01	0	-0.005	6.84E-01
LDL-C adjusted for HDL &TG	MVMR	70	0.012 (-0.35,0.37)	9.49E-01					
TG	IVW	46	-0.001 (-0.19,0.19)	9.91E-01	9.91E-01				
TG	MR Egger	46	0.08 (-0.39,0.54)	7.38E-01		1.00E+00	0	-0.0055	6.61E-01
TG adjusted for HDL &LDL	MVMR	46	0.035 (-0.22,0.29)	7.88E-01					

For Abbreviations, see Table 3.3.

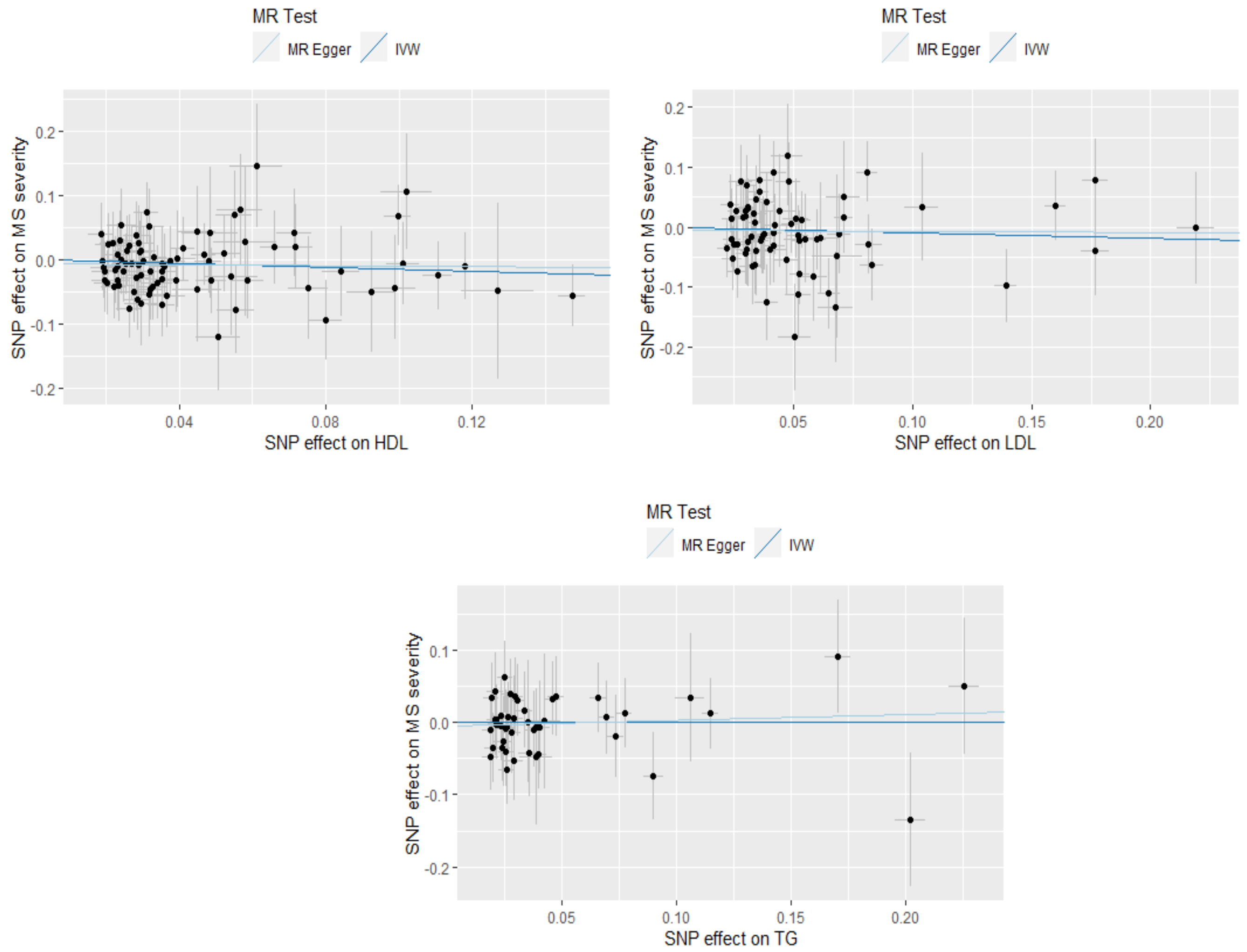


Figure 3.5: Scatter plots for MR analyses showing the causal estimates of the lipid fractions on MS severity. The effect sizes of each genetic variant (with 95% confidence intervals) are represented by black points. The slope of each line shows the estimated MR effect for each method.

3.3.2. Genetically predicted MS risk not associated with lipid levels (reverse causation analysis)

I sought to explore whether the liability to MS risk would exert a change in lipid levels. To do so, I initially selected 200 autosomal susceptibility SNPs outside the MHC region that reported by the IMSGC as genome-wide significant for MS (methodology chapter section 2.2.3). These SNPs account for almost 19% of the MS heritability. The selected SNPs clumped at an LD threshold value of $r^2 < 0.01$. With MS-associated SNPs as the exposure, I obtained corresponding effect estimates for HDL-C, LDL-C and TG from GLGC (methodology chapter section 2.2.2.2) as the outcome.

The remaining 118 and 119 SNPs out of the 200 were subjected to further MR analysis. The mean F-statistics of these SNPs was around 75, suggesting that bias due to weak instruments is negligible in this data. The IVW and MR-Egger results revealed no causal link between the genetic determinants of MS risk and HDL-C, LDL-C or TG (**Table 3.6** and **Figure 3.6**). The associations between genetically predicted MS risk and the lipid fractions are presented as 1-SD with 95% CI per 1-unit-higher log-odds of MS risk. There was evidence of significant heterogeneity; however, the MR-Egger intercept test suggested no evidence of pleiotropy.

Table 3.6: MR analysis of the effect of MS risk on lipid levels

Outcome	Method	No. of SNP	Beta (95 % CI)	p-value	FDR	Pleiotropy assessment			
						Q p-value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p-value
HDL-C	IVW	118	-0.004 (-0.02,0.01)	5.91E-01	8.86E-01				
HDL-C	MR Egger	118	-0.016 (-0.06,0.03)	4.98E-01		4.05E-12	54	0.0012	5.93E-01
LDL-C	IVW	118	-0.008 (-0.02,0.01)	2.95E-01	8.85E-01				
LDL-C	MR Egger	118	0 (-0.04,0.05)	9.91E-01		2.61E-07	44.5	-0.0008	7.17E-01
TG	IVW	119	0 (-0.01,0.01)	9.61E-01	9.61E-01				
TG	MR Egger	119	0.017 (-0.02,0.06)	4.05E-01		2.55E-06	41.6	-0.0017	3.70E-01

For Abbreviations, see Table 3.3.

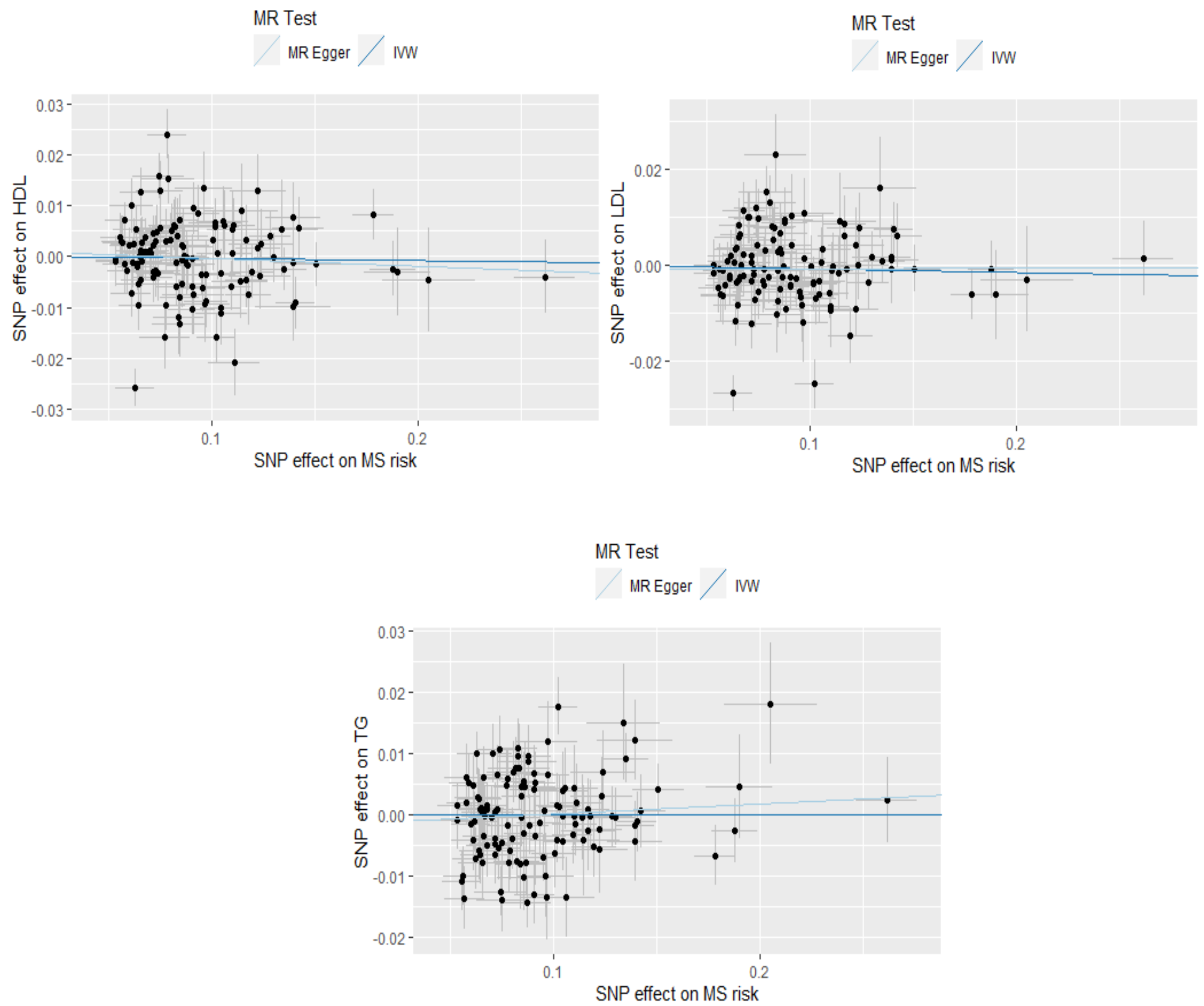


Figure 3.6: Scatter plots for MR analyses showing the causal estimates of the MS risk on lipid fractions. The effect sizes of each genetic variant (with 95% confidence intervals) are represented by black points. The slope of each line shows the estimated MR effect for each method.

3.3.3. MR estimates for mimicking the effects of statins using expression data

To investigate the potential role of and mechanisms used by statins in MS risk and severity, I used eQTL data from the eQTLGen (methodology chapter section 2.2.2.2) to genetically mimic statin effects. I used whole-blood *cis*-eQTL in a ±5 kilobases flank around 25 genes (including *HMGCR*) that encode proteins involved in cholesterol biosynthesis to genetically mimic the effects statins elicit via the cholesterol-dependent pathway (**Table 3.7**). I also used whole-blood *cis*-eQTL in a ±5 kilobases flank around 20 Rho GTPase gene regions to genetically mimic the effects statins elicit via cholesterol-independent pathways (**Table 3.7**). The primary focus was on Rho GTPases that undergo prenylation, a total of 16 family members (**Table 3.7**). Narrow flanking regions were used to reduce the possibility of the selected SNPs associating with MS via pathways other than those involving the target genes. All the selected SNPs were associated with the genes at p-values $< 5 \times 10^{-8}$ and clumped at the liberal LD-clumping threshold value of $r^2 < 0.4$. This liberal LD-clumping threshold was selected because it was difficult to find many independent SNPs in a single gene region. Next, corresponding effect estimates for these SNPs on MS risk and severity were obtained, and IVW and MR-Egger methods were implemented, considering the correlation between the genetic instruments (methodology chapter section 2.3.2.5). Since the selected SNPs were in LD (correlated), the strength of the instruments could not be measured because the R^2 and F-statistics tests required independent instruments. However, MR assumes that *cis*-acting SNPs are robust instruments and ideal for MR analysis because they strongly affect a gene’s expression level and are more likely to satisfy the key assumptions for MR (Porcu et al., 2019).

Table 3.7: The Rho GTPase family and mevalonate pathway (cholesterol biosynthesis pathway) gene list (*part 1*)

HGNC symbol	Ensemble Gene ID	Description	Participant of	Genes - MS risk analyses	Genes - MS severity analyses
PPAPDC2	ENSG00000205808	phosphatidic acid phosphatase type 2 domain containing 2	Cholesterol biosynthesis	Yes	Yes
HMGCS1	ENSG00000112972	3-hydroxy-3-methylglutaryl-CoA synthase 1 (soluble)	Cholesterol biosynthesis	Yes	Yes
TM7SF2	ENSG00000149809	transmembrane 7 superfamily member 2	Cholesterol biosynthesis	Yes	Yes
ARV1	ENSG00000173409	ARV1 homolog (S. cerevisiae)	Cholesterol biosynthesis	Yes	Yes
IDI1	ENSG00000067064	isopentenyl-diphosphate delta isomerase 1	Cholesterol biosynthesis	Yes	Yes
HMGCR	ENSG00000113161	3-hydroxy-3-methylglutaryl-CoA reductase	Cholesterol biosynthesis	Yes	Yes
LBR	ENSG00000143815	lamin B receptor	Cholesterol biosynthesis	Yes	Yes
PMVK	ENSG00000163344	phosphomevalonate kinase	Cholesterol biosynthesis	Yes	Yes
FDFT1	ENSG00000079459	farnesyl-diphosphate farnesyltransferase 1	Cholesterol biosynthesis	Yes	Yes
DHCR24	ENSG00000116133	24-dehydrocholesterol reductase	Cholesterol biosynthesis	Yes	Yes
MVK	ENSG00000110921	mevalonate kinase	Cholesterol biosynthesis	Yes	NO
FDPS	ENSG00000160752	farnesyl diphosphate synthase	Cholesterol biosynthesis	Yes	Yes
ACAT2	ENSG00000120437	acetyl-CoA acetyltransferase 2	Cholesterol biosynthesis	Yes	Yes
GGPS1	ENSG00000152904	geranylgeranyl diphosphate synthase 1	Cholesterol biosynthesis	Yes	Yes
CYP51A1	ENSG00000001630	cytochrome P450, family 51, subfamily A, polypeptide 1	Cholesterol biosynthesis	Yes	Yes
HSD17B7	ENSG00000132196	hydroxysteroid (17-beta) dehydrogenase 7	Cholesterol biosynthesis	Yes	Yes
MVD	ENSG00000167508	mevalonate (diphospho) decarboxylase	Cholesterol biosynthesis	Yes	Yes
SQLE	ENSG00000104549	squalene epoxidase	Cholesterol biosynthesis	Yes	Yes
SC5DL	ENSG00000109929	sterol-C5-desaturase	Cholesterol biosynthesis	Yes	Yes
LSS	ENSG00000160285	lanosterol synthase (2,3-oxidosqualene-lanosterol cyclase)	Cholesterol biosynthesis	Yes	Yes
MSMO1	ENSG00000052802	methylsterol monooxygenase 1	Cholesterol biosynthesis	NO	NO
EBP	ENSG00000147155	emopamil binding protein (sterol isomerase)	Cholesterol biosynthesis	NO	NO
IDI2	ENSG00000148377	isopentenyl-diphosphate delta isomerase 2	Cholesterol biosynthesis	NO	NO
NSDHL	ENSG00000147383	NAD(P) dependent steroid dehydrogenase-like	Cholesterol biosynthesis	NO	NO
DHCR7	ENSG00000172893	7-dehydrocholesterol reductase	Cholesterol biosynthesis	Yes	NO
RHOV	ENSG00000104140	ras homolog family member V	member of Rho GTPase	NO	NO
RHOB	ENSG00000143878	ras homolog family member B	member of Rho GTPase	Yes	Yes
CDC42	ENSG00000070831	cell division cycle 42	member of Rho GTPase	Yes	Yes
RHOG	ENSG00000177105	ras homolog family member G	member of Rho GTPase	Yes	Yes
RHOD	ENSG00000173156	ras homolog family member D	member of Rho GTPase	Yes	NO
RHOF	ENSG00000139725	ras homolog family member F (in filopodia)	member of Rho GTPase	Yes	Yes
RHOQ	ENSG00000119729	ras homolog family member Q	member of Rho GTPase	Yes	Yes
RAC2	ENSG00000128340	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	member of Rho GTPase	Yes	Yes
RAC1	ENSG00000136238	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	member of Rho GTPase	Yes	Yes
RHOH	ENSG00000168421	ras homolog family member H	member of Rho GTPase	Yes	Yes
RHOBTB2	ENSG00000008853	Rho-related BTB domain containing 2	member of Rho GTPase	Yes	Yes

Table 3.7: The Rho GTPase family and mevalonate pathway (cholesterol biosynthesis pathway) gene list (*part 2*)

HGNC symbol	Ensemble Gene ID	Description	Participant of	Genes - MS risk analyses	Genes - MS severity analyses
RAC3	ENSG00000169750	ras-related C3 botulinum toxin substrate 3 (rho family, small GTP binding protein Rac3)	member of Rho GTPase	NO	NO
RHOU	ENSG00000116574	ras homolog family member U	member of Rho GTPase	Yes	NO
RHOBTB1	ENSG00000072422	Rho-related BTB domain containing 1	member of Rho GTPase	Yes	Yes
RND2	ENSG00000108830	Rho family GTPase 2	member of Rho GTPase	NO	NO
RHOC	ENSG00000155366	ras homolog family member C	member of Rho GTPase	Yes	Yes
RND1	ENSG00000172602	Rho family GTPase 1	member of Rho GTPase	Yes	Yes
RND3	ENSG00000115963	Rho family GTPase 3	member of Rho GTPase	NO	NO
RHOA	ENSG00000067560	ras homolog family member A	member of Rho GTPase	NO	NO
RHOJ	ENSG00000126785	ras homolog family member J	member of Rho GTPase	NO	NO

The table includes 25 genes that were flagged as being involved in the cholesterol biosynthesis pathway in the Reactome database of human pathways and reactions (http://www.reactome.org). The 20 genes of the Rho GTPase family were extracted from (Azzarelli et al., 2015). For each gene, the HGNC symbol (HUGO Gene Nomenclature Committee), Ensemble gene ID and description were derived from the Ensembles database (https://www.ensembl.org/index.html). In the last two columns, 'Yes' indicates that a gene was included in the MR analysis, while 'NO' indicates that a gene was not included in the MR analysis due to either no SNPs being robustly associated with the target gene at a p-value $< 5 \times 10^{-8}$ or the eQTL data being absent.

3.3.3.1. Genetically mimicked effect of statins on MS risk is independent of cholesterol

Table 3.8 displays the associations between the genetically mimicked statin effects and MS risk (ORs with 95% CI per 1-SD higher expression of the target gene in blood). A total of 35 genes (21/25 genes of the cholesterol biosynthesis pathway and 14/20 genes of the Rho GTPase family) were selected for analysis on the basis of having at least one genetic variant strongly associated with their expression. MR analyses involving SNPs in these gene regions found only a link between the expression levels of *RAC2* and MS risk that was significant (FDR ≤ 0.05) after multiple testing corrections.

The heterogeneity, in general, in these analyses ranged from non-significant to moderate, and the MR-Egger intercept test provided no evidence for horizontal pleiotropy except for *RHOH*.

For *RAC2,* the IVW result revealed that one SD increase in genetically predicted *RAC2* expression in the blood was associated with a 14% reduction in MS risk (**Table 3.8** and **Figure 3.7**). The MR-Egger causal estimate was significant and largely consistent with the IVW results, reducing the probability that pleiotropy influenced these results. There was no evidence for heterogeneity, and the MR-Egger intercept test provided no evidence for directional pleiotropy. Since the results were deemed significant after multiple testing corrections, replication was assessed using the whole-blood *cis*-eQTL dataset from the GTEx project (methodology chapter section 2.2.2.2). It was found that the direction of the effect was identical across the discovery and replication results, providing further support for *RAC2* playing a protective role in MS risk (**Table 3.8**).

Table 3.8: MR estimates for the genetically mimicked effects of statins on MS risk (*part 1*)

Trait	Method	No. of SNPs	OR (95 % CI)	p-value	FDR	pleiotropy assessment			
						Q p-value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p-value
ACAT2	IVW	5	1.035(0.92,1.17)	5.77E-01	8.74E-01				
ACAT2	MR Egger	5	1.043(0.89,1.22)	6.00E-01		6.27E-01	0	-0.0076	4.85E-01
ARV1	IVW	10	0.939(0.86,1.03)	1.84E-01	7.38E-01				
ARV1	MR Egger	10	0.941(0.84,1.05)	2.96E-01		6.79E-01	0	-4.00E-04	9.52E-01
CYP51A1	IVW	3	1.029(0.89,1.19)	6.99E-01	8.74E-01				
CYP51A1	MR Egger	3	1.046(0.38,2.85)	9.30E-01		7.93E-01	0	-0.0036	9.69E-01
DHCR24	IVW	9	1.053(0.92,1.2)	4.37E-01	8.74E-01				
DHCR24	MR Egger	9	1.024(0.73,1.44)	8.90E-01		9.61E-01	0	0.0072	8.12E-01
DHCR7	Wald ratio	1	0.431(0.18,1.02)	5.68E-02	6.05E-01				
FDFT1	IVW	25	0.984(0.95,1.02)	4.07E-01	8.74E-01				
FDFT1	MR Egger	25	0.979(0.94,1.02)	2.62E-01		3.33E-01	0.05 3	0.0058	6.12E-02
FDPS	Wald ratio	1	0.878(0.57,1.35)	5.53E-01	8.74E-01				
GGPS1	IVW	6	1.009(0.9,1.13)	8.74E-01	8.74E-01				
GGPS1	MR Egger	6	1.018(0.84,1.24)	8.61E-01		3.29E-01	0	-0.0027	9.16E-01
HMGCR	IVW	3	1.178(0.89,1.56)	2.50E-01	7.38E-01				
HMGCR	MR Egger	3	1.074(0.09,12.56)	9.55E-01		3.33E-01	0	0.0101	9.41E-01
HMGCS1	Wald ratio	1	1.877(0.92,3.82)	8.24E-02	6.05E-01				
HSD17B7	IVW	9	1.061(0.83,1.36)	6.40E-01	8.74E-01				
HSD17B7	MR Egger	9	1.132(0.76,1.69)	5.45E-01		8.56E-02	0.36	-0.0101	7.01E-01
IDI1	IVW	4	0.941(0.59,1.51)	8.00E-01	8.74E-01				
IDI1	MR Egger	4	0.493(0.21,1.17)	1.10E-01		4.97E-02	0.5	0.0491	1.13E-01
LBR	IVW	3	0.77(0.48,1.22)	2.69E-01	7.38E-01				
LBR	MR Egger	3	0.608(0,985.61)	8.95E-01		8.75E-01	0	0.0174	9.46E-01
LSS	IVW	15	0.993(0.94,1.05)	8.03E-01	8.74E-01				
LSS	MR Egger	15	1.006(0.93,1.09)	8.82E-01		3.67E-02	0.4	-0.0028	6.84E-01
MVD	IVW	8	1.061(0.76,1.47)	7.26E-01	8.74E-01				
MVD	MR Egger	8	1.047(0.73,1.5)	8.01E-01		7.61E-03	0.6	0.0019	9.16E-01
MVK	IVW	3	1.079(0.78,1.49)	6.39E-01	8.74E-01				
MVK	MR Egger	3	1.237(0.14,10.68)	8.47E-01		7.63E-01	0	-0.0148	8.87E-01

Table 3.8: MR estimates for the genetically mimicked effects of statins on MS risk (*part 2*)

Trait	Method	No. of SNPs	OR (95 % CI)	p-value	FDR	pleiotropy assessment			
						Q p-value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p-value
PMVK	IVW	4	0.877(0.69,1.11)	2.81E-01	7.38E-01				
PMVK	MR Egger	4	0.741(0.38,1.43)	3.73E-01		1.85E-01	0.11	0.0248	5.91E-01
PPAPDC2	IVW	9	1.065(0.97,1.17)	2.02E-01	7.38E-01				
SC5DL	IVW	4	0.989(0.89,1.1)	8.43E-01	8.74E-01				
SC5DL	MR Egger	4	0.993(0.82,1.2)	9.40E-01		3.21E-01	0	-0.0013	9.62E-01
SQLE	IVW	5	1.024(0.82,1.28)	8.38E-01	8.74E-01				
SQLE	MR Egger	5	1.054(0.73,1.52)	7.80E-01		6.65E-01	0	-0.0067	7.69E-01
TM7SF2	IVW	3	0.877(0.75,1.02)	8.64E-02	6.05E-01				
TM7SF2	MR Egger	3	1.068(0.37,3.06)	9.03E-01		8.01E-01	0	-0.0489	6.56E-01
PPAPDC2	MR Egger	9	0.961(0.83,1.11)	5.86E-01		3.01E-01	0.045	0.0321	7.91E-02
RAC1	IVW	10	1.071(0.96,1.2)	2.41E-01	5.41E-01				
RAC1	MR Egger	10	1.218(0.9,1.66)	2.08E-01		7.11E-01	0	-0.0168	3.56E-01
RAC2	IVW	15	0.861(0.78,0.95)	**3.80E-03**	**5.00E-02**				
RAC2	MR Egger	15	0.855(0.76,0.96)	**8.14E-03**		6.73E-01	0	0.0033	5.15E-01
RHOB	IVW	3	0.849(0.53,1.36)	4.94E-01	5.83E-01				
RHOB	MR Egger	3	1.053(0.07,15.89)	9.70E-01		6.53E-01	0	-0.0172	8.66E-01
RHOBTB1	IVW	13	0.977(0.81,1.18)	8.09E-01	8.09E-01				
RHOBTB1	MR Egger	13	1.017(0.86,1.21)	8.48E-01		4.66E-02	0.4	-0.0154	2.42E-01
RHOBTB2	IVW	3	0.713(0.51,1)	4.94E-02	2.21E-01				
RHOBTB2	MR Egger	3	0.868(0.36,2.11)	7.56E-01		5.70E-01	0	-0.0273	5.61E-01
RHOC	IVW	5	1.038(0.83,1.3)	7.44E-01	8.02E-01				
RHOC	MR Egger	5	1.167(0.75,1.82)	4.94E-01		5.00E-02	0.49	-0.0268	5.62E-01
RHOD	IVW	2	0.544(0.29,1.03)	6.30E-02	2.21E-01				
RHOF	IVW	6	1.134(0.93,1.39)	2.18E-01	5.41E-01				
RHOF	MR Egger	6	0.873(0.43,1.79)	7.12E-01		6.80E-01	0	0.0523	4.43E-01
RHOG	IVW	4	1.126(0.85,1.49)	4.07E-01	5.69E-01				
RHOG	MR Egger	4	1.228(0.84,1.8)	2.91E-01		2.38E-01	0	-0.0153	4.97E-01
RHOH	IVW	5	0.518(0.29,0.94)	3.01E-02	2.11E-01				
RHOH	MR Egger	5	1.733(0.74,4.07)	2.06E-01		4.14E-01	0	-0.0905	**2.41E-03**
RHOQ	IVW	9	1.07(0.88,1.3)	4.99E-01	5.83E-01				
RHOQ	MR Egger	9	0.909(0.66,1.25)	5.56E-01		5.79E-01	0	0.0261	1.70E-01
RHOU	IVW	3	1.161(0.89,1.51)	2.70E-01	5.41E-01				
RHOU	MR Egger	3	1.327(0.49,3.61)	5.79E-01		6.08E-01	0	-0.0196	7.59E-01
RND1	Wald ratio	1	0.805(0.5,1.3)	3.77E-01	5.69E-01				
CDC42	IVW	10	1.054(0.93,1.19)	3.98E-01	5.69E-01				
CDC42	MR Egger	10	1.069(0.86,1.32)	5.43E-01		3.83E-02	0.45	-0.0041	8.85E-01
Replication									
RAC2	IVW	2	0.70(0.51,0.96)	**2.80E-02**					

Genes highlighted with orange encode proteins involved in cholesterol biosynthesis, while genes highlighted with green encode members of the Rho family. Pleiotropy assessment cannot be conducted for instruments consisting of ≤ 2 independent SNPs as it requires > 2 SNPs. *For Abbreviations, see Table 3.3.*

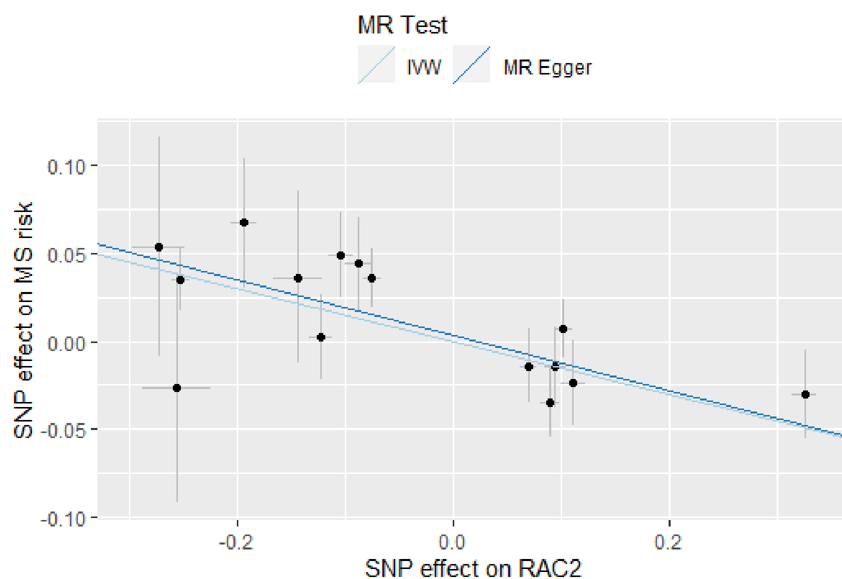


Figure 3.7: Scatter plots for MR analyses showing the causal estimates of *RAC2* on MS risk. The effect sizes of each genetic variant (with 95% confidence intervals) are represented by black points. The slope of each line shows the estimated MR effect for each method.

3.3.3.2. Genetically mimicked effect of statins had no causal association with MS severity

Table 3.9 displays the associations between the genetically mimicked statin effects and MS severity (log ORs with the 95% CI per 1-SD higher expression of the target gene in blood). A total of 31 genes (19/25 genes involved in the cholesterol biosynthesis pathway and 12/20 genes of the Rho GTPase family) were selected for analysis on the basis of having at least one genetic variant strongly associated with their expression. The MR results showed no evidence of an association between the SNPs in these genes and MS severity. There was no evidence for heterogeneity or horizontal pleiotropy in these MR analyses.

Table 3.9: MR estimates for the genetically mimicked effects of statins on MS severity (*part 1*)

Trait	Method	No. of SNPs	Beta (95 % CI)	p-value	FDR	pleiotropy assessment			
						Q p-value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p-value
ACAT2	IVW	3	-0.145(-0.55,0.26)	4.86E-01	9.36E-01				
ACAT2	MR Egger	3	-0.111(-0.58,0.36)	6.42E-01		3.75E-01	0	-0.0129	4.60E-01
ARV1	IVW	3	0.034(-0.24,0.31)	8.08E-01	9.51E-01				
ARV1	MR Egger	3	0.234(-0.46,0.93)	5.09E-01		4.75E-01	0	-0.0656	4.98E-01
CYP51A1	Wald ratio	1	-0.155(-0.63,0.32)	5.24E-01	9.36E-01				
DHCR24	IVW	2	0.149(-0.24,0.54)	4.56E-01	9.36E-01				
FDFT1	IVW	8	0.163(0,0.33)	5.41E-02	9.36E-01				
FDFT1	MR Egger	8	0.151(-0.04,0.34)	1.17E-01		1.02E-01	0.34	0.0042	8.34E-01
FDPS	Wald ratio	1	0.152(-1.22,1.52)	8.28E-01	9.51E-01				
GGPS1	IVW	2	0.003(-0.32,0.33)	9.88E-01	9.88E-01				
HMGCR	Wald ratio	1	0.274(-0.65,1.2)	5.62E-01	9.36E-01				
HMGCS1	Wald ratio	1	-1.142(-3.19,0.9)	2.73E-01	9.36E-01				
HSD17B7	Wald ratio	1	0.163(-1.21,1.54)	8.17E-01	9.51E-01				
IDI1	IVW	3	-0.11(-0.94,0.72)	7.95E-01	9.51E-01				
IDI1	MR Egger	3	6.549(0.91,12.19)	2.29E-02		2.57E-01	0	-0.3974	1.92E-02
LBR	IVW	3	-0.399(-1.74,0.94)	5.58E-01	9.36E-01				
LBR	MR Egger	3	-0.129(-5.54,5.28)	9.63E-01		5.17E-01	0	-0.0196	9.15E-01
LSS	IVW	5	-0.133(-0.28,0.01)	7.15E-02	9.36E-01				
LSS	MR Egger	5	-0.134(-0.29,0.02)	9.42E-02		3.93E-01	0	0.0007	9.80E-01
MVD	Wald ratio	1	0.227(-0.94,1.39)	7.02E-01	9.47E-01				
PMVK	Wald ratio	1	0.482(-0.6,1.56)	3.81E-01	9.36E-01				
PPAPDC2	IVW	3	0.105(-0.14,0.35)	4.04E-01	9.36E-01				
PPAPDC2	MR Egger	3	0.212(-12.8,13.23)	9.75E-01		9.69E-01	0	-0.0279	9.85E-01
SC5DL	IVW	2	0.073(-0.26,0.4)	6.65E-01	9.47E-01				
SQLE	Wald ratio	1	0.038(-1.41,1.48)	9.59E-01	9.88E-01				
TM7SF2	IVW	2	-0.135(-0.65,0.38)	6.04E-01	9.36E-01				
CDC42	IVW	2	0.012(-0.73,0.75)	9.75E-01	9.88E-01				
RAC1	IVW	3	-0.398(-0.92,0.12)	1.34E-01	9.36E-01				
RAC1	MR Egger	3	-0.765(-3.44,1.91)	5.75E-01		5.45E-01	0	0.0631	7.76E-01
RAC2	IVW	5	0.095(-0.25,0.44)	5.88E-01	9.36E-01				
RAC2	MR Egger	5	0.099(-0.59,0.79)	7.79E-01		7.90E-01	0	-0.0012	9.85E-01
RHOB	Wald ratio	1	-0.646(-2.79,1.5)	5.55E-01	9.36E-01				
RHOBTB1	IVW	5	0.347(-0.15,0.84)	1.70E-01	9.36E-01				
RHOBTB1	MR Egger	5	0.325(-0.65,1.3)	5.12E-01		2.78E-01	0	0.0042	9.56E-01
RHOBTB2	IVW	2	-0.467(-1.78,0.85)	4.86E-01	9.36E-01				
RHOC	IVW	2	0.211(-0.45,0.87)	5.31E-01	9.36E-01				
RHOF	IVW	2	0.113(-0.44,0.67)	6.92E-01	9.47E-01	3.75E-01			4.60E-01

Table 3.9: MR estimates for the genetically mimicked effects of statins on MS severity (*part 2*)

Trait	Method	No. of SNPs	Beta (95 % CI)	p-value	FDR	pleiotropy assessment			
						Q p-value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p-value
RHOG	IVW	4	-0.366 (-1.12,0.39)	3.39E-01	9.36E-01				
RHOG	MR Egger	4	-0.231 (-2.1,1.63)	8.09E-01		8.03E-01	0	-0.023	7.79E-01
RHOH	IVW	2	-0.85 (-2.25,0.55)	2.35E-01	9.36E-01				
RHOQ	IVW	4	-0.317 (-0.99,0.36)	3.58E-01	9.36E-01				
RHOQ	MR Egger	4	-0.27 (-1.68,1.14)	7.08E-01		3.57E-01	0	-0.0054	9.40E-01
RND1	Wald ratio	1	0.212 (-2.14,2.57)	8.60E-01	9.52E-01				

Genes highlighted with orange encode proteins involved in cholesterol biosynthesis, while genes highlighted with green encode members of the Rho family. Pleiotropy assessment cannot be conducted for instruments consisting of ≤ 2 independent SNPs as it requires > 2 SNPs.

For Abbreviations, see Table 3.3.

3.4. Discussion

The work presented in this chapter aimed to: (1) dissect the causal nature of the association between blood lipid levels and MS (risk and severity) and explore whether genetic predisposition to increased major plasma lipid fractions plays an aetiological role in MS; (2) explore the potential effects of statins on MS via MR analysis conducted using SNPs in different gene regions that genetically mimic statin biological effects; and (3) assess whether there is reverse causation between lipid fractions and MS risk.

3.4.1. High plasma HDL-C is a risk factor for MS

Recent epidemiological findings indicated that there is a correlation between dyslipidaemia (elevated plasma LDL-C and TGs and low plasma HDL-C) and MS disease activity and disability progression (Weinstock-Guttman et al., 2011, Tettey et al., 2014, Zhornitsky et al., 2016). However, it is unclear if and how plasma lipid levels and functions are altered in patients with MS and whether such changes influence disease development and severity or if they are only useful as biomarkers of disease activity. Several MR analyses have been conducted in the present study to address these questions. The results show that lifelong high HDL-C leads to a high MS risk. This finding is reproducible and robust in terms of heterogeneity, pleiotropy and reverse causation testing. In contrast, genetically raised circulating TGs are unlikely to be associated with the risk of developing MS.

Associations between lipids and MS risk have received insufficient attention in epidemiological studies. Surprisingly, only one MR analysis on lipids and MS risk with GLGC and IMSGC data, the same datasets used in the current study, has been published (Yuan et al., 2021). The primary findings of that study demonstrated that there is no causal role for genetically raised LDL-C and TGs on MS risk, and there was only weak evidence of association between genetically raised HDL-C and MS risk (IVW OR = 1.14, p-value = 0.057) (Yuan et al., 2021).

The MR results of the current study agree with those of that study regarding LDL-C and TGs but not HDL-C—I found robust evidence of an HDL-C–MS risk causal association. The most notable difference is the number of SNPs included in the analysis model, which may explain why previous results differ from current results regarding HDL-C. In the aforementioned study, 68 SNPs were used to genetically proxy circulating levels of HDL-C, and they explained about 1.6% of the variance in HDL-C levels. In the current study, I used 118 SNPs to genetically proxy circulating levels of HDL-C, and they explained about 9% of the variance in HDL-C levels, clearly more than the variances explained by the 68 SNPs in the previous MR study. Thus, the MR model used here had sufficient power to detect a causal association between HDL-C and MS risk.

Elevated plasma HDL-C levels are reported in patients with MS at different stages of disease, including RR and clinical inactivity (Gafson et al., 2018, Rádiková et al., 2020, Jorissen et al., 2017). Although the mechanisms responsible for the increased risk of MS in patients with elevated HDL-C are unclear, the current and previously reported findings highlight the potential role of circulating HDL-C levels in MS development and its clinical course.

It is well-known that raising HDL-C levels significantly decreases the risk of CHD. This beneficial effect is attributed to HDL-C's key role in reverse cholesterol transport, which results in its anti-atherogenic, anti-oxidant and anti-inflammatory properties (Kosmas et al., 2018). However, HDL-C may not always be protective and can become dysfunctional. Several lines of non-genetic evidence have shown that HDL-C in patients at increased risk of some conditions can lose its anti-inflammatory and anti-oxidant properties and become dysfunctional (Ansell et al., 2003). For example, in acute and chronic inflammation conditions (e.g. type 2 diabetes and atherosclerosis), replacing the anti-inflammatory apolipoprotein A-I (apoA-I) with the inflammatory serum amyloid A protein (SAA) can convert HDL particles from anti-inflammatory to pro-inflammatory, which inherently limits HDL-C's protective features (Rao and Kakkar, 2011, Ansell et al., 2005).

Furthermore, genetic studies of lipid levels have revealed genes that affect circulating HDL-C levels (Weissglas-Volkov and Pajukanta, 2010). These genes include variants ranging from common variants that have minor effects to rare mutations that cause complete loss of the gene's activity, which may increase susceptibility to diseases (Weissglas-Volkov and Pajukanta, 2010). Finally, it is worth mentioning that statins can cause modest increases in HDL-C levels among statin users (McTaggart and Jones, 2008, Ansell et al., 2003), an effect independent of LDL-C reduction (Barter et al., 2010). This HDL-C increase achieved by statin appears to enhance the anti-inflammatory ability of HDL-C (Ansell et al., 2003). In the current study, statins were not expected to contribute to the causal association between HDL-C and MS risk since individuals who were known to be on lipid-lowering medication were excluded from the GLGC data (Willer et al., 2013).

3.4.2. The lower cholesterol levels induced by statins have no effect on MS risk

In the second part of this work, I conducted a separate MR analysis to address the causal link between genetically mimicked statin effects and MS. First, I used variants in *HMGCR* and other downstream genes to mimic the cholesterol-dependent effects of statins in relation to MS risk. The findings suggest that stains do not reduce MS risk through mechanisms that contribute to cholesterol level reduction. This result was expected, because LDL-C itself does not have a causal role in MS risk in the current results. Therefore, the current findings suggest that if statins effectively reduce MS risk, the effects would be mediated via a pathway that is independent of circulating cholesterol reduction. This finding, in line with a recent study, suggests that the beneficial effects of simvastatin in patients with MS are independent of serum cholesterol (Eshaghi et al., 2019). In that study, the authors reanalysed the phase 2 MS-STAT trial data obtained from 140 patients with secondary progressive MS who were randomised to receive placebo or simvastatin; the data consisted of brain magnetic resonance imaging (MRI) reports, physical disability (EDSS) scores and serum total cholesterol levels that were recorded at baseline and after one and two years (Eshaghi et al., 2019). They applied structural equation models to examine whether the beneficial effects of simvastatin on reducing the rate of brain atrophy and slowing deterioration are dependent on or independent of blood cholesterol reduction (Eshaghi et al.,

2019). Their results showed that the cholesterol-independent model significantly explained simvastatin's effect on the clinical and MRI outcome measures (Eshaghi et al., 2019).

3.4.3. The effects induced by statins via the cholesterol-independent pathway (*RAC2*) may reduce MS risk

Since the lowering of cholesterol showed no effect on MS risk, my attention was directed to exploring the causal link between Rho GTPases and MS risk. Interestingly, the MR results showed that genetically predicted *RAC2* expression was causally associated with reducing MS risk, and this finding emerged as robust with sensitivity analysis and was replicated in an independent eQTL dataset.

RAC2 is a Rho GTPase family member (**Table 3.1**) expressed mainly in blood cell lineages, indicating that it has specific cellular functions in these cells (Tell et al., 2012). *RAC2* regulates multiple key processes of inflammatory responses, including dendritic cell migration, nicotinamide adenine dinucleotide phosphatase (NADPH) oxidase activity and T-cell proliferation, migration and differentiation to the Th1 subtype (Saoudi et al., 2014, Sironi et al., 2011). In addition to immune activation, *RAC2* is involved in the induction of peripheral immune tolerance. It is an essential component of restimulation-induced cell death (Ramaswamy et al., 2007), a necessary process in the self-limiting negative feedback mechanism used to control T-cell expansion during ongoing immune responses (Fattouh et al., 2013). Furthermore, *RAC2* modulates atherosclerotic calcification by regulating IL-1β production by macrophages (Ceneri et al., 2017).

The impact of *RAC2* on MS pathology has not yet been confirmed; however, an association between *RAC2* and MS has previously been reported. It has been shown that *RAC2* haplotypes are associated with susceptibility to MS and earlier onset of disease symptoms (Sironi et al., 2011). Furthermore, *RAC2* levels in whole blood samples from patients with MS were found to be low compared to those in healthy controls due to upregulation of *RAC2* microRNAs (Yang et al., 2017).

MicroRNAs are small non-coding RNAs that negatively regulate the expression of their target genes (Yang et al., 2017).

Like other Rho GTPases, Rac2 protein undergoes isoprenylation and cycles between an inactive GDP-bound state and an active GTP-bound state to regulate a variety of cellular processes (Bayo et al., 2021). Recent findings suggest that the *RAC2* represents a pleiotropic effect of statin therapy. It has been shown that statins, through inhibition of isoprenylation of Rac2, reduce oxidative stress during sepsis and downregulate pentraxin 3 in vascular cells during immune-inflammatory responses (Delbosc et al., 2002, Durant et al., 2004, Habib et al., 2007, Baetta et al., 2015). Furthermore, statins have been shown to induce the expression of several genes, including *RAC2,* that are involved in epidermal growth factor signalling (Sawaya et al., 2019); however, the mechanism by which statins can induce *RAC2* expression remains to be identified. Therefore, it is reasonable to suggest that statins may reduce the risk of MS through a mechanism that involves *RAC2*.

Taken together, the current results shed light on the role *RAC2* plays as a potent genetic modifier of MS risk. In addition, it can be postulated that statins might mediate some beneficial effects on MS risk via RAC2-regulated pathways. Nonetheless, caution should be taken to avoid overinterpretation of these findings. Although MR is a powerful tool for investigating the causal relationship between an exposure and an outcome, this approach cannot identify the underlying mechanism of the relationship or confirm the hypothesis in the current study regarding statins, *RAC2* and MS risk. In addition, the possibility that *RAC2* reducing the risk of MS is independent of statins effect cannot be ruled out. Thus, further studies are required to identify the mechanism responsible for the observed causal relationship between *RAC2* and MS risk and to test the hypothesis that statins reduce MS risk via a RAC2-related mechanism.

3.4.4. Plasma lipids have no causal link to MS severity

This study also highlights that genetic susceptibility to high levels of circulating lipid fractions does not affect MS severity. Despite several epidemiological studies investigating the associations between circulating lipid fractions and accrual of disability in patients with MS, most of these studies used EDSS to measure the disability and a few used MS severity scores. The difference between these measures is that the MS severity score has better metric properties that correct the EDSS for disease duration (Weinstock-Guttman et al., 2011). The current MR findings support those of a previous retrospective cohort study that involved 3,166 participants with MS and found no statistically significant association between hyperlipidaemia and EDSS (Zhang et al., 2018). The current findings are also in line with those of another study conducted with 1,083 patients with RR in which comorbidities were identified through a database linked to electronic medical records. In that study, it was observed that hyperlipidaemia had no significant effect on walking speed and self-reported disability (Conway et al., 2017). These two studies accounted for different covariates, such as sex, age, cohort entry year, use of disease-modifying drugs and socioeconomic status (Conway et al., 2017, Zhang et al., 2018). However, the current MR findings are contrary to those previously reported. For example, a retrospective study that included 492 patients with MS showed that worsening EDSS and MS severity was associated with higher baseline LDL-C and TG but not HDL-C (Weinstock-Guttman et al., 2011). This study did not account for potential confounders, apart from sex and age. In another cohort of 178 participants with MS, while no association was observed between HDL-C and EDSS and MS severity, higher LDL-C and TG levels were significantly associated with higher EDSS and MS severity (Tettey et al., 2014). The association between LDL-C and TG levels and EDSS diminished after accounting for age, sex, BMI and physical activity but remained significant for MS severity (Tettey et al., 2014). Indeed, confounding and reverse causality in observational studies cannot be entirely ruled out. In addition, these studies used small sample sizes to measure the change in disability within two-year periods, whereas longer follow-up and large sample sizes are preferred when measuring the change in disability.

In the current study, MR approach was used, which limited the potential bias associated with the presence of confounders.

3.4.5. Genetically mimicked effects of statins have no causal links to MS severity

Finally, as with MS risk, the MR approach was implemented to address the causal role of variants in the gene regions that mimic the cholesterol-dependent and cholesterol-independent pathways in MS severity. No evidence of association was found between these variants and MS severity, indicating that statins might not reduce MS severity via mechanisms associated with the genes targeted in the current study.

To the best of my knowledge, the impact of statin treatment on disability progression measured by the MS severity score has not yet been studied. A handful of studies have explored the impact of statins on disability progression measured by the EDSS; however, the results were inconclusive. A recent phase 2 MS-STAT trial that included 140 patients with MS found that those using simvastatin (80 mg per day) had a small change in EDSS score after two years, but it was significant (Chataway et al., 2014). In contrast, an open-label non-placebo-controlled trial study found that lovastatin treatment (40 mg per day) over one year had no effect on the EDSS score of seven participants with MS (Sena et al., 2003). A separate retrospective review of disability progression in patients with MS that included 45 patients with MS treated with statins and 90 control patients with MS who did not receive statin therapy found no difference in EDSS score between the groups, at least at statin doses currently prescribed for hyperlipidaemia (Soldán et al., 2012). The possible explanation for this apparent contradiction is that the phase 2 MS-STAT trial had a larger sample size and the statin doses were larger than the doses in the latter two studies, indicating the possibility that higher doses of statins may effective to reduce the worsening of disability in patients with MS.

Although the genetic findings in the current study agree with the above studies that statins do not affect disability progression, it is important to note that eQTL data were measured in blood not brain tissue; therefore, it is possible that genetically mimicked statins effect may be missed out in this work .It has been hypothesised that various mechanisms contribute to the accumulation of disability in progressive MS, including chronic microglial activation, mitochondrial dysfunction, compartmentalisation of inflammation behind an intact BBB and formation of ectopic lymphoid tissue in the meninges.

3.4.6. limitations

The main advantage of the MR approach is that it overcomes the common biases (reverse causality and confounding) found in conventional observational studies owing to the random allocation of alleles at conception. A range of applicable methods were adopted here in both the causality estimation and sensitivity analyses to yield reliable results and assess the robustness of the MR estimates against potential violations, specifically, the direction of pleiotropy.

However, one limitation of the current study is that the major lipid fractions (HDL-C, LDL-C and TG) are each heterogeneous groups of particles defined by differences in particle size, density, apoprotein content, migration characteristics and relationships to disease, and these subfractions differ in their risk profiles (Rádiková et al., 2020). For example, increases in cardiovascular risk, metabolic syndrome and type 2 diabetes have been found to be associated with small dense LDL-C particles, while no association has been observed between large LDL-C particles and cardiovascular risk. Moreover, large HDL-C particles may be more atheroprotective than small HDL-C particles (Rádiková et al., 2020). This study was designed to investigate total blood lipid levels and thus did not consider whether there are subtypes of these fractions (e.g. LDL sub-particles) (Zhornitsky et al., 2016, Rádiková et al., 2020) that might play different roles in MS risk or severity.

3.4.7. Conclusion

Taken together, evidence from this study supports the existence of a causal effect of HDL-C on MS risk and shows that there is no reverse causation; however, no evidence for the causal role of LDL-C and TGs on MS risk was found. The MR findings suggest that *RAC2* is a potent genetic modifier of MS risk. Since it has been reported to mediate some of the pleiotropic effect of statins, it is possible to postulate that statins may mediate some of their pleiotropic effect on MS risk via *RAC2* . However, further studies are required to either support or reject this hypothesis. No evidence was found of a causal effect for lipid-related traits on MS severity. Finally, the current genetic evidence did not support the repurposing of statin treatment to reduce severity.

Chapter 4 *Applying Mendelian randomization to identify potential druggable targets for multiple sclerosis*

Statement of contribution

I executed the analysis scripts used in this chapter myself. Scripts used to generate the analysis provided by Catherine Strom. The druggable genome list provided by Aroon Hingorani and Chris Finan. Nicholas Wood and Catherine Strom contributed to the interpretation of my results.

4.1. Introduction

Thus far, there has been no cure for MS, and the primary goal of licenced DMTs is to modify the course of the disease, reduce the frequency and duration of relapses, control symptoms and slow disability. However, some DMTs are hampered by potentially serious adverse reactions and require careful monitoring through a specialist MS clinic (**Table 1.4, Chapter one**) (De Angelis et al., 2018). Newer DMTs have shown better short-term outcomes than older DMTs, but the available data on their long-term effectiveness and harm are insufficient (De Angelis et al., 2018). The costs of these therapies are also a major issue, particularly in low- and middle-income countries. At present, many MS DMTs cost beyond US$90,000 a year, and their economic value has been widely debated (Hartung, 2021). These reports increase the need for more research to discover new and better interventions and preventive measures for MS. Although drug discovery is viewed as a vital endeavour to improve the treatment and management of disorders, discovering new drugs for human diseases is a lengthy, complex and costly process with high attrition rates (on Neuroscience and Disorders, 2014, Baird et al., 2021, Reay and Cairns, 2021). These challenges facing the pharmaceutical industry have prompted an increasing focus on drug repurposing, also known as drug repositioning (Reay and Cairns, 2021). Drug repurposing is an alternative strategy that involves identifying novel indications for existing approved drugs or drugs in clinical trials (Cha et al., 2018). This strategy has the potential to complement traditional drug discovery by mitigating the high monetary- and time-related costs and risks linked with the latter (Cha et al., 2018). Interestingly, recent estimates have shown that almost 30% of drugs approved by the US Food and Drug Administration were repurposed (Rudrapal et al., 2020).

The rapidly growing body of GWASs, whole-genome studies and whole-exome sequencing studies have contributed remarkably to the rapid progress in identifying thousands of variants within genetic loci that influence human health and disease and provide evidence to detect molecular pathways involved in disease (Nelson et al., 2015, Gaziano et al., 2021, Baird et al., 2021). Indeed, genetic

insights increase the odds of success in drug discovery and repurposing; it has been shown that drug targets with human genetic evidence of disease association are more than twice as likely to lead to approved drugs compared to those without (Gaziano et al., 2021, Baird et al., 2021, Reay and Cairns, 2021, King et al., 2019).

One approach to facilitate drug repurposing through genetics is using MR to integrate GWAS associations with QTL to infer whether a causal relationship exists between protein targets for existing drugs and disease indications. One interesting example of MR-based drug repurposing is Parkinson's disease. Storm et al. performed drug repurposing MR between expression and protein QTL (measured in the blood and brain) and Parkinson's disease and found 23 drug-targeting mechanisms for Parkinson's disease, including four possible drug repurposing opportunities and two drugs that may increase Parkinson's disease risk (Storm et al., 2021).

A conventional MR analysis uses GWAS SNPs associated only with exposure as instrumental variables to estimate the causal effects of these instruments on an outcome through the exposure of interest. These GWAS SNPs may not be substantial enough to inform drug repurposing, as the vast majority are common SNPs that have a small effect size on the phenotype of interest (Reay and Cairns, 2021). Therefore, QTLs are preferred in drug repurposing MR. QTL SNPs exert their impact on molecular phenotypes, such as gene expression; therefore, their effect size is large enough to be a relevant proxy for treatment (Reay and Cairns, 2021). In addition, *cis*-acting QTL SNPs on druggable protein levels or gene expressions that encode druggable proteins are viewed as powerful tools for informing therapeutic targeting because they mimic the on-target (beneficial or harmful) effects observed by pharmacological agents (Gaziano et al., 2021). The direction of the relationship between expression and the trait can also be exploited to inform drug repurposing; for example, if the increased genetically predicted expression of a particular gene is associated with a disease risk or progression, then an antagonist of that gene may be useful, and vice versa (Reay and Cairns, 2021).

4.2. Aims

I sought to perform the druggable genome study using an MR approach to identify and prioritise new drug targets to treat MS's disease. I employed MR on gene expression measured in blood and brain tissue to identify druggable genes whose (genetically predicted) expression is causally related to MS risk or MS severity.

4.3. Method

A summary of the methods and workflow of MR in this study are reported in Figure 4.1. To examine the causal links between the expression levels of druggable genes and MS, a list of curated druggable genes was first obtained from the corresponding authors, Professor Aroon Hingorani and Dr Chris Finan (Finan et al., 2017). Finan and colleagues identified 4,863 druggable genes classified into four tiers:

- Tier 1—genes encoding targets of approved and clinical-phase drugs (Storm et al., 2021).
- Tier 2—genes not targeted by existing drugs but encode proteins targeted by small, drug-like molecules *in vivo* and genes encoding proteins with high sequence homology to approved drug targets (Storm et al., 2021).
- Tiers 3 and 4—genes encoding proteins with more distant similarity to approved drug targets, secreted or extracellular proteins and members of key druggable gene families not already included in Tiers 1 and 2. Of these, genes in Tier 3 were located within 50 kb of a GWAS SNP and encoded an extracellular protein. The remainder were placed in Tier 4 (Storm et al., 2021).

Then, I obtained the genetic instruments for each druggable gene from blood and brain-tissue-specific *cis*-eQTL from eQTLGen and PsychENCODE consortia (methodology section 2.2.2.2). SNP outcome associations were obtained from the discovery phase for MS risk (the MHC region was excluded) and MS severity (see Section 2.2.3). SNPs strongly associated with the genes of interest at p-value 5×10^{-8} and located within 5 kilobases around each gene were kept.

SNP exposure and outcomes were harmonised, and palindromic SNPs were removed. SNPs were then clumped at a clumping threshold value of $r^2 < 0.2$ and then filtered out using the Steiger filtering method. Next, the selected SNPs were included in the MR model to obtain the overall effects of the estimates using IVW as the main estimator and MR-Egger as a sensitivity estimator. Due to the LD correlation among SNPs, the extended IVW and MR-Egger were used to account for SNPs'

correlations by incorporating the LD matrix. If only one SNP was available for exposure, the Wald ratio was used to estimate the MR effect. As additional sensitivity analyses, Cochran's Q and I^2 tests were used to assess heterogeneity and the MR-Egger intercept to assess pleiotropy. If the Cochran's Q and/or MR-Egger intercept p-value was < 0.05, a scatter plot was used as a visual tool to identify potential outliers. These outliers were removed, and the analysis was repeated with the remaining SNPs.

P-values were adjusted for multiple tests using FDR ≤ 0.05, adjusting for the number of genes analysed per exposure–outcome combination.

All genes that reached significance (FDR ≤ 0.05) using the Wald ratio or IVW method were replicated in an independent MS risk case-control cohort (14,498 subjects with MS and 24,091 healthy controls) (Beecham et al., 2013a).

4.4. Results

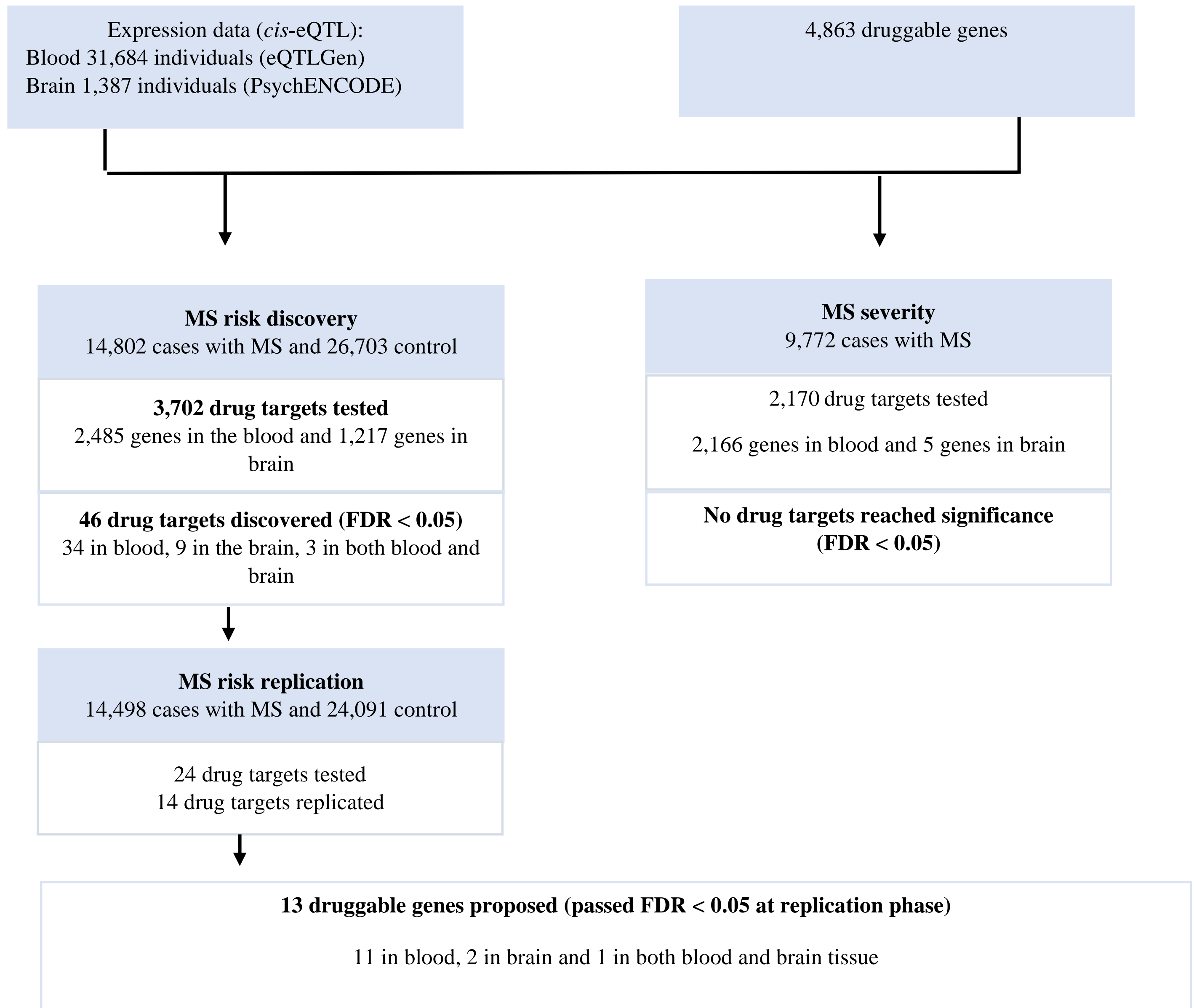


Figure 4.1: Workflow and summarised results of the current study.

4.4.1. MS's disease risk – discovery and replication

The genetically determined expression of 46 genes was significantly associated with MS risk in the discovery cohort at FDR ≤ 0.05. Of these 46 genes, 34 were in blood tissue, 9 were in brain tissue, and 3 were in both blood and brain tissues (**Table A.1**).

I took the 46 genes forward to replication in an independent MS disease case-control cohort. The genetically predicted expression of 14 genes was replicated: 11 in blood tissue, 2 in brain tissue and 1 in both blood and brain (**Figure 4.2 and Table A.1**). Of these 14 genes, *CCR4, CD6, IFNGR2, IL7, MAPK3, MAST3, SIK3, SLAMF7, STAT3, TNFSF14, TYMP, GALC* and *IL2RA* reached FDR ≤ 0.05, while *CDK14* only reached nominal significance (p-value ≤ 0.05).

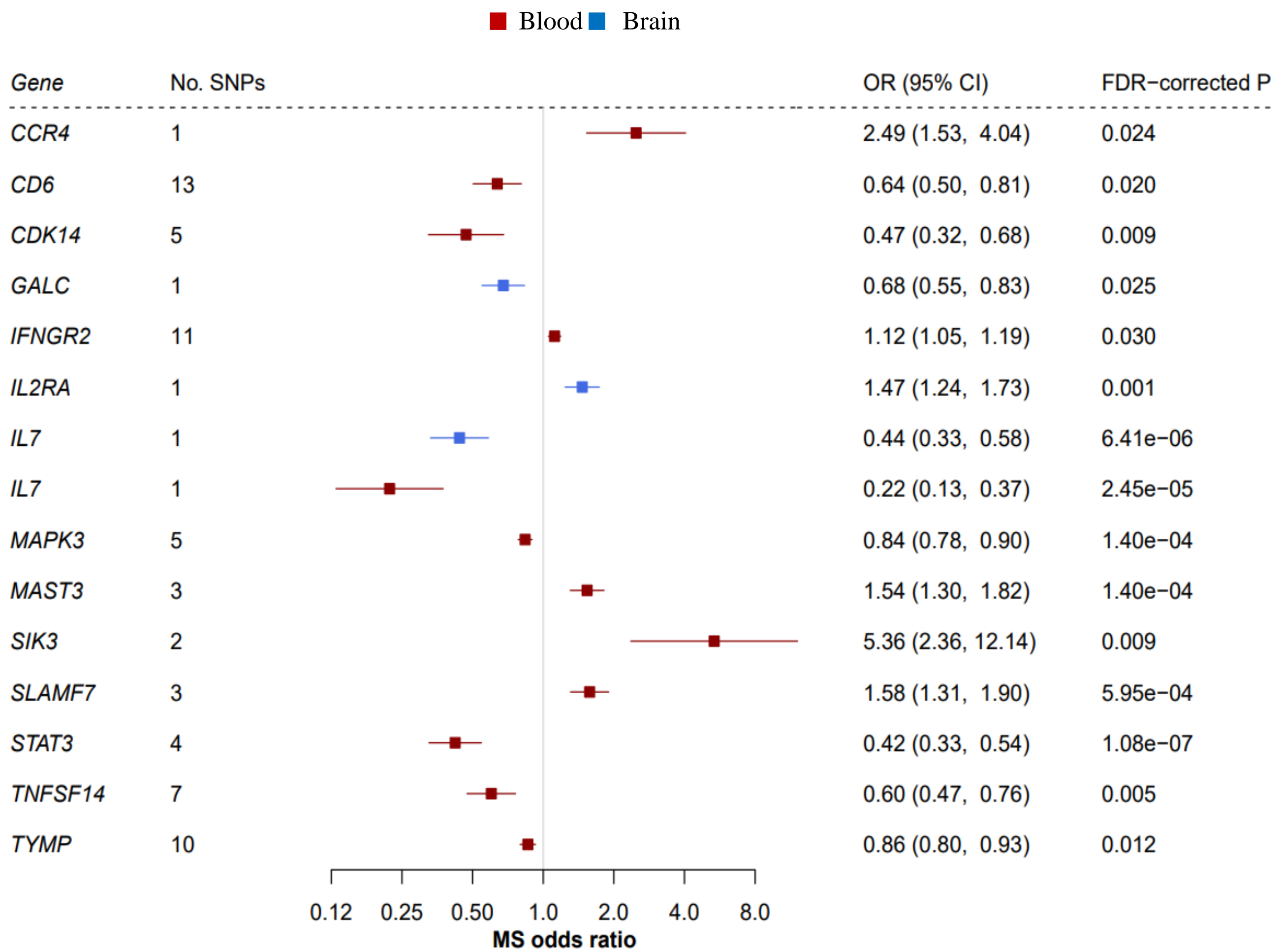


Figure 4.2: The forest plot visualises the discovery phase results for the 14 replicated genes. The results from the Wald ratio (if the number of SNPs < 2) or inverse variance-weighted method (if the number of SNPs ≥ 2) are displayed. Each point represents MS's disease odds ratio per one standard deviation increase in gene expression. Horizontal lines represent the 95% confidence interval. Colour codes: red = blood, blue = brain tissue.

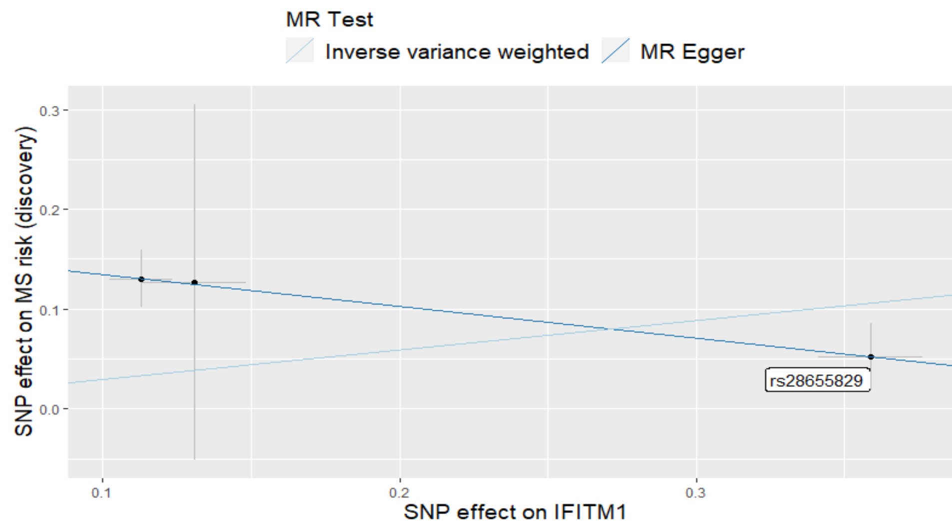
Overall, in the discovery phase, the magnitude and direction of MR estimates were largely consistent between IVW and MR-Egger for all MR analyses except for *IFITM1, ITGB3* and *SAE1* (**Table A.1**). For these three genes, the direction of the estimates from MR-Egger was opposite to the direction of IVW estimates, indicating the existence of invalid SNPs. The heterogeneity and pleiotropy tests for *ITGB3* were significant (Cochran's Q and MR-Egger intercept p-value < 0.05) but not for *IFITM1* and *SAE1*. Perhaps Cochran's Q and MR-Egger intercept p-values had low power due to the small quantity

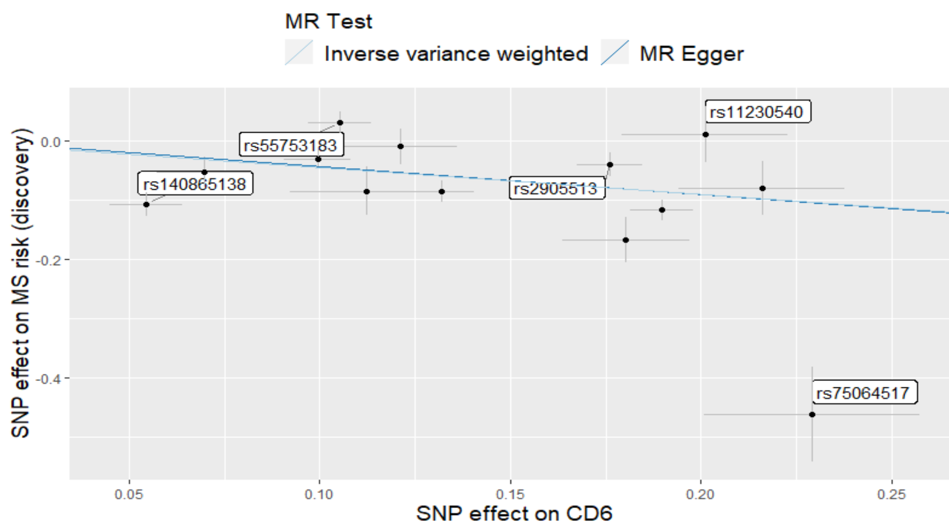
of SNPs. The statistical power for Cochran's Q increased when the number of SNPs increased, whereas the power of the MR-Egger intercept increased when there was variability in the instrument strengths across the set of SNPs (Bowden et al., 2017, Bowden et al., 2018a). Additionally, three more genes in the discovery phase did not pass the heterogeneity test, including *CD6*, *TNFSF14* and *TYMP* (**Table A.3**). No evidence of pleiotropy was detected for the remaining genes (**Table A.3**).

To identify potential outlier (invalid) SNPs, scatter plots were used to detect the outliers for the genes that did not pass the sensitivity tests (Cochran's Q and MR-Egger intercept tests) to obtain more reliable results (**Figure 4.2**). These outliers were removed, and MR analyses were repeated. Visual inspection of the scatter plots revealed five outliers for *CD6*, two outliers for *TNFSF14* and one outlier for each of *ITGB3, TYMP, SAE1* and *IFITM1.* These outliers were removed one by one until the heterogeneity and/or pleiotropy were eliminated. After the outliers' removal, the direction of estimates became consistent among the MR methods, and the Cochran's Q and MR-Egger intercepts became non-significant. The IVW estimates obtained by excluding outliers remained statistically significant (**Table 4.1**).

In the replication phase, MR estimates' magnitude and direction were consistent across MR methods, the exception being *STAT3* (**Table A.1**). All the replicated genes, except *CD6*, passed the pleiotropy and heterogeneity tests (**Table A.3**). Scatter plots showed one outlier for *CD6* and *STAT3* (**Figure 4.3**). For *STAT3,* after removing the outlier, the direction of the estimates turned out to be consistent between IVW and MR-Egger. For *CD6*, the MR-Egger intercept turned out to be non-significant after removing the outlier. The IVW estimates obtained by excluding outliers remained statistically significant for both *CD6* and *STAT3*.

The direction of MR estimates was the same in both the discovery and replication cohorts for the 13 gene that passed the significance in the replication phase (**Table A.1**).





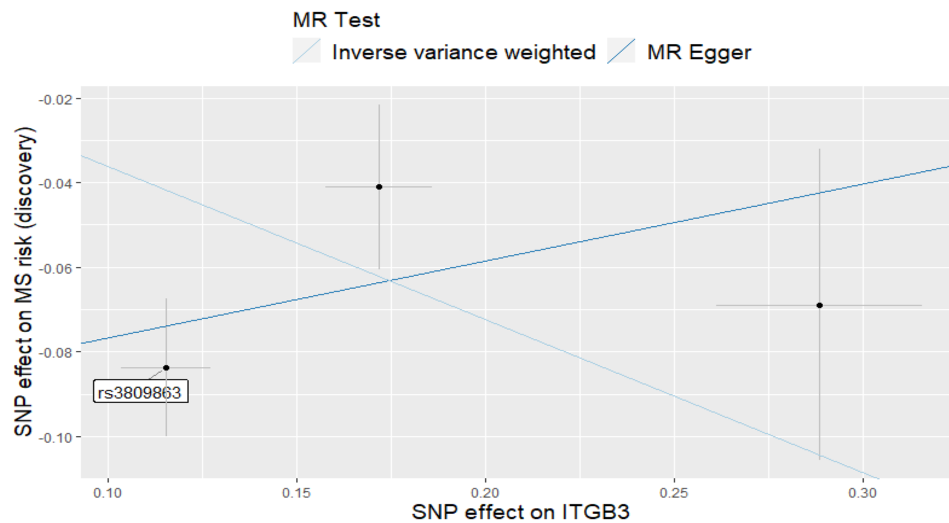


Figure 4.3: Scatter plots of causal estimates of druggable genes that did not pass the sensitivity tests for MS risk (part 1). Outliers are labelled.

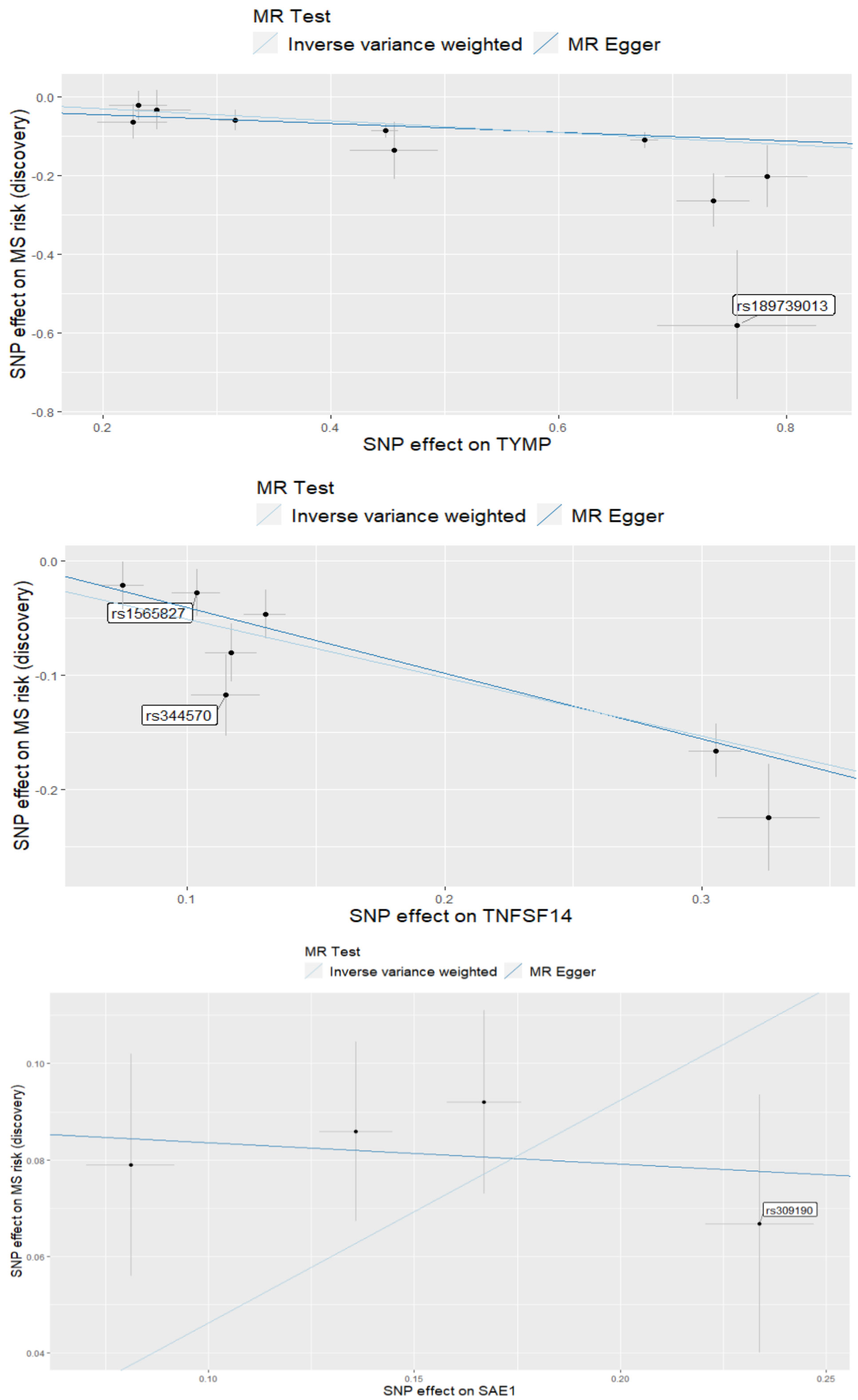


Figure 4.3: Scatter plots of causal estimates of druggable genes that did not pass the sensitivity tests for MS risk (part 2). Outliers are labelled.

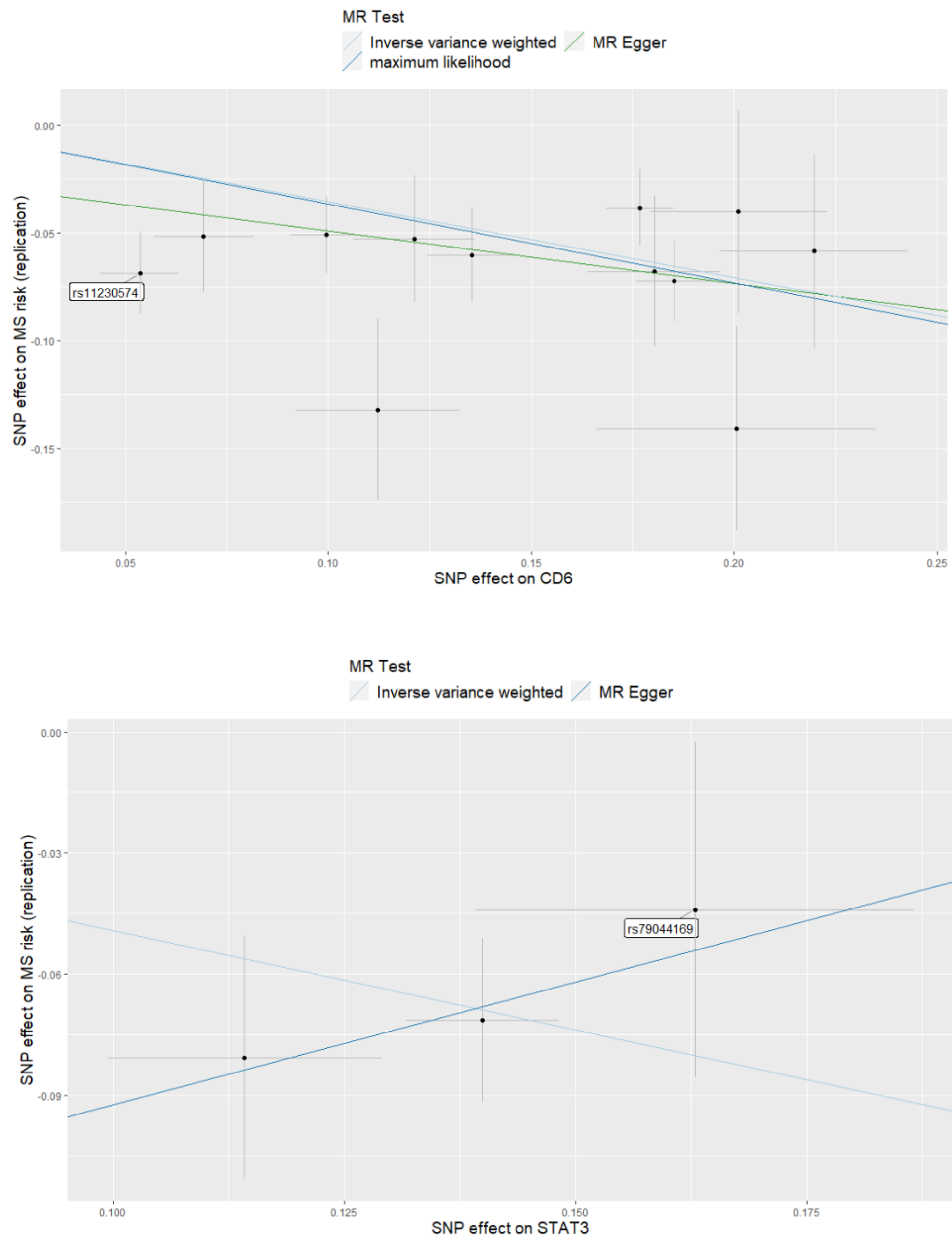


Figure 4.3: Scatter plots of causal estimates of druggable genes that did not pass the sensitivity tests for MS risk (part 3). Outliers are labelled.

Table 4.1: MR results on effects of druggable genes on MS risk after excluding outliers

Gene	Druggability tier	Outcome	No. SNP	Method	Beta	SE	p-value	Tissue	Cochrans' Q p-value	I^2	MR-Egger intercept p-value
ITGB3	1	MS risk (discovery)	2	IVW	-0.239	0.08	4.73E-03	Blood			
IFITM1	3	MS risk (discovery)	2	IVW	1.15	0.25	4.56E-06	Blood			
CD6	1	MS risk (discovery)	8	IVW	-0.591	0.09	6.34E-12	Blood			
CD6	1	MS risk (discovery)	8	MR-Egger	-0.53	0.19	4.51E-03	Blood	8.82E-02	0.36	7.22E-01
SAE1	2	MS risk (discovery)	3	IVW	0.584	0.1	1.19E-08	Blood			
SAE1	2	MS risk (discovery)	3	MR-Egger	0.153	3.98	9.69E-01	Blood	9.42E-01	0	9.08E-01
TNFSF14	1	MS risk (discovery)	5	IVW	-0.547	0.09	4.05E-09	Blood			
TNFSF14	1	MS risk (discovery)	5	MR-Egger	-0.598	0.12	1.42E-06	Blood	1.42E-01	0.27	5.51E-01
TYMP	2	MS risk (discovery)	9	IVW	-0.164	0.03	3.71E-09	Blood			
TYMP	2	MS risk (discovery)	9	MR-Egger	-0.132	0.06	1.70E-02	Blood	2.15E-01	0.16	5.11E-01
STAT3	1	MS risk (replication)	2	IVW	-0.556	0.127	1.28E-05	Blood			
CD6	1	MS risk (replication)	11	IVW	-0.39	0.076	2.89E-07	Blood			
CD6	1	MS risk (replication)	11	MR-Egger	-0.076	0.203	7.07E-01	Blood	2.50E-01	0.12	1.03E-01

4.4.2. MS severity

Of the druggable genes, 2,170 had an eQTL available for MR using MS severity data, of which 2,166 genes were tested in blood, and only 8 genes were tested in brain tissue (**Table A.2**). Of the 2,166 genes, only 125 (120 genes in blood and 4 genes in brain and one gene in both blood and brain tissue) reached nominal significance (p-value $< 5 \times 10^{-8}$) but did not pass significance after multiple testing corrections (FDR ≤ 0.05). None of the tested genes in the brain reached nominal significance (**Table A.4**).

4.4.3. Discussion

This study sought to combine MS GWAS data with an eQTL dataset and a list of curated druggable genes to identify new druggable targets for MS risk and MS severity. Genetic evidence in favour of 13 prioritised drug targets for MS risk was provided. These were prioritised through using the MR approach, and they emerged as robust based on sensitivity analyses. They were further confirmed using an independent MS risk cohort. Six druggable genes out of 13 showed genetic evidence of a causal relationship between the genetically predicted expression of *CCR4, IFNGR2, IL2RA, MAST3*, *SIK3* and *SLAMF7* and increased MS risk. The remaining seven genes showed evidence of protection that genetically predicted the expressions of *CD6, GALC, IL7, MAPK3, STAT3, TNFSF14* and *TYMP,* which were casually associated with reduced MS risk.

Three replicated genes (*CCR4, SLAMF7* and *SIK3*) that encode protein targets for existing drugs as priority candidates were identified for evaluation in randomised trials of early management for MS. These drugs are either approved medications or in a clinical trial phase, and their pharmacological effects are consistent with the direction of MR effect estimates for the above genes (**Figure 4.2**).

CCR4 is a protein-coding gene belonging to the CC chemokine receptor family, a family of chemotactic cytokines known for their roles in leukocyte activation and chemotaxis (Ness et al., 2006). *CCR4* is expressed on dendritic cells, macrophages, NK cells, platelets and basophils, but it has been shown to be predominantly expressed on T-cells, especially on Th2 and regulatory T-cells (Watson and Marx, 2019, Ness et al., 2006). Upon binding to its ligands, CCL17 and CCL22, *CCR4* promotes the cell trafficking of lymphocytes to various organs, including the skin (Watson and Marx, 2019). *CCR4* is a potential target for mogamulizumab, an approved medication for mycosis fungoides and Sezary syndrome. T-cells migrate to the skin during these conditions due to the combined effect of *CCR4* overexpression on malignant T-cells and the overproduction of *CCR4* ligands by various skin cells (Watson and Marx, 2019). Mogamulizumab selectively binds to *CCR4* to mark T-cells for destruction through an antibody-dependent cellular cytotoxicity process (Watson and

Marx, 2019). Drug safety evaluations have identified mogamulizumab as an effective and well-tolerated therapy for patients with fungoides and Sezary syndrome (Afifi et al., 2019).

CCR4 has been reported as one of the risk genes for MS (Consortium*† et al., 2019, Beecham et al., 2013a), and several clinical studies of MS, together with experimental studies in animals, have demonstrated the functional roles of CCR4-CCL17/CCL22 in the pathogenesis of MS and EAE. For example, studies have identified the CCL22 protein in CNS-infiltrating leukocytes and the microglia of EAE-induced mice, and *CCR4* is expressed by invading leukocyte subsets (Poppensieker et al., 2012). Shimizu et al. found that the numbers of $CD4^{+}CCR4^{+}$ cells in the peripheral blood of patients with MS were higher in the relapsing phase than in the remission phase or in healthy controls (Shimizu et al., 2011). Furthermore, evidence from EAE models of MS observed a delay and decreased disease incidence in CCR4-deficient mice with diminished pathogenic T-cell filtration to the CNS (Forde et al., 2011). These results support the pathogenic role of *CCR4* in MS and further suggest *CCR4* antagonists as a possible strategy for the prevention of MS.

Although there is no ongoing clinical trial for *CCR4* and MS, a few studies have attempted to examine *CCR4* antagonist effects on CNS autoimmunity using EAE models; however, the results were inconsistent. One study found that compound 22, a *CCR4* antagonist, significantly ameliorated EAE, suggesting that *CCR4* antagonism is a potential therapeutic strategy for MS (Scheu et al., 2017). However, another study found no evidence of the effects of the *CCR4* antagonist AF399/420/18025 on the EAE clinical score (Scheu et al., 2017). It has been argued that the absence of the effect in the latter study is due to the use of dimethyl sulfoxide as a drug vehicle solution for *CCR4* antagonists, which may affect the BBB and thereby influence the infiltration of mononuclear cells into the CNS (Scheu et al., 2017).

SLAMF7 encodes a protein member of the SLAM family of receptors, which have roles in cytotoxicity, humoral immunity, autoimmunity, cell survival, cell adhesion, and lymphocyte development (Ritchie and Colonna, 2018). It was found to be highly expressed on malignant plasma cells and normally expressed on leukocyte subsets, such as NK cells, $CD8^{+}$ T-cells, activated monocytes and dendritic cells (Ritchie and Colonna, 2018).

Elotuzumab is a humanised immunoglobulin G monoclonal antibody targeting *SLAMF7*, which is in clinical use for the treatment of multiple myeloma (Ritchie and Colonna, 2018). Elotuzumab induces myeloma cell death via two mechanisms: (1) by binding to *SLAMF7* on myeloma cells and promoting interactions with NK cells to mediate the killing of myeloma cells via an antibody-dependent cellular cytotoxicity mechanism and (2) by directly binding to *SLAMF7* on natural killer cells and causing costimulatory signalling that further potentiates the killing of myeloma (Ritchie and Colonna, 2018).

GWAS prioritises *SLAMF7* as MS susceptibility loci on chromosome one (Beecham et al., 2013a). The role and function of *SLAMF7* in MS pathogenesis have not been established. However, *SLAMF7* contributes to highly activated macrophage-driven inflammation in autoimmune and infectious diseases; its activation signature has been shown to be up-regulated in inflammatory diseases, such as rheumatoid arthritis, inflammatory bowel disease and COVID-19 pneumonia (Simmons et al., 2022). The current results showed that the genetically predicted expression of *SLAMF7* was associated with increased MS risk, suggesting that targeting *SLAMF7* may be useful in preventing MS.

SIK3 is a member of the SIKs, a subfamily belonging to the AMP-activated protein kinase family. SIKs are ubiquitously expressed in humans, and SIK3 is predominantly expressed in neural tissues (Sun et al., 2020). SIKs are involved in controlling gene expression in response to extracellular cues that increase intracellular cAMP levels (Jin et al., 2020). Recent evidence has indicated that SIKs play a role in modulating the production of pro- and anti-inflammatory cytokines in myeloid cells, such as macrophages and dendritic cells. For example, it has been shown that SIK inhibition via pharmacological inhibition compounds, such as HG-9-91-01, enhances the macrophage's anti-inflammatory phenotype, which is characterised by the production of high levels of anti-inflammatory cytokines, such as IL-10, and low levels of pro-inflammatory cytokines, including tumour necrosis factor α and IL-6 (Darling et al., 2017, Jin et al., 2020). In addition, evidence suggests that *SIK3* plays an important role in T-cells.

Proteomic studies on peripheral T-cells revealed that *SIK3* is expressed in T-cells and strongly up-regulated following TCR stimulation , suggesting an important role(s) for SIK3 in T-cells (Nefla et al., 2021).

Tenalisib (RP6530) is a small molecule inhibitor of PI3K delta and gamma isoforms and *SIK3*, which is currently in phase 2 clinical development for locally advanced or metastatic breast cancer. The mechanism of action for tenalisib on *SKI3* is unclear, but data show that this agent can inhibit the PI3K delta and gamma isoforms and prevent the stimulation of the PI3K/AKT-mediated signalling pathway. This may result in a reduction in cellular proliferation in PI3K delta/gamma-expressing tumour cells. Furthermore, tenalisib is suggested as a modulator for inflammatory responses via several mechanisms, for instance, by preventing the release of reactive oxygen species (ROS) from neutrophils and tumour necrosis factor (TNF)-alpha activity. Interestingly, ROS has been implicated as a mediator of demyelination and axonal damage in both MS and EAE, and TNF-alpha is a major cytokine that plays a pivotal role in the pathogenesis of MS (Gilgun-Sherki et al., 2004, Zahid et al., 2021). Together, these data, alongside the current results, underscore the potential of *SIK3* as a druggable target for MS that is worth further investigation.

This study found that two of the replicated genes are already being targeted by existing compounds for MS treatment: *IL2RA* and *IL-7* (**Table 4.2**). IL2RA is a target for daclizumab, a previously approved medication for relapsing forms of MS with a pharmacological effect consistent with the direction of the MR effect estimate in the current study. However, in 2018, daclizumab was withdrawn from the market following reports of serious and potentially fatal immune reactions affecting the brain (including encephalitis and meningoencephalitis), liver, and other organs (https://www.gov.uk/).

Whereas *IL-7* is a target for GSK2618960, a molecule that was under investigation as a treatment for MS (identifier: NCT01808482), its trial was terminated at phase 1 due to the misrepresentation of preclinical data that supported the rationale for GSK2618960 in MS, as stated in the *Clinical trials.go* database. GSK2618960 is a humanised IgG1 monoclonal antibody that functions as an antagonist, competitively inhibiting IL-7 binding and blocking IL-7 signalling via IL-7Rα (Liao et al., 2021). Regardless of the termination of the study, the GSK2618960 mechanism of action is not consistent with the direction of MR effect estimates in the current study, that is, genetically predicted *IL-7* expression associated with reduced MS risk (**Figure 4.2**).

IL-7 is a vital cytokine that promotes T-cell development, survival and proliferation by engaging with its receptor IL-7Rα, thus contributing to the homeostasis of the peripheral T-cell pool (Mazzucchelli et al., 2012, Lei et al., 2017). Therefore, it is not surprising that a growing body of evidence suggests the role of IL-7−IL-7Rα signalling dysregulation in the pathogenesis of MS (Liao et al., 2021).

Genetic evidence for the *IL-7* pathway in the pathogenesis of MS identified a nonsynonymous missense SNP, rs6897932, located in the 6th exon of the *IL-7Rα* gene (Vandenbroeck, 2012). This mutation leads to an increase in the levels of the soluble form of IL-7Rα (Vandenbroeck, 2012). The SNP rs6897932 was associated with MS susceptibility in the MS GWAS in 2011(Sawcer et al., 2011). In a recent MS GWAS in 2019 (Consortium*† et al., 2019), the *IL-7Rα* gene was also nominated as a susceptibility locus for MS and showed a new SNP, rs11567694, which was in the complete LD with rs6897932 (D' and r^2 values of 1.0). Individuals with genotype-induced elevations of soluble IL7Rα have been demonstrated to show an increased risk of autoimmune disease, including MS, due to the potentiation of IL-7 bioactivity (Lundström et al., 2013).

This is because soluble IL-7Rα competes with cell-associated IL-7 receptors for binding to IL-7 molecule (Lundström et al., 2013). Under normal circumstances, IL-7 is a limited resource whose level is regulated mainly via receptor-mediated clearance, engaging the soluble IL-7Rα with IL-7 molecules, resulting in less IL-7 clearance and ultimately increased IL-7 bioavailability and bioactivity (Lundström et al., 2013). The presence of soluble IL-7Rα contributes to IL-7 bioactivity by enhancing IL-7, which induces T-cell survival, augmenting the proliferative responses of T-cells to weak self-antigens and diminishing Fas-induced cell death (Lundström et al., 2013). Given this, IL-7/IL-7Rα pathway agonism may ameliorate MS, whereas antagonism may enable IL-7 to bind with soluble IL-7Rα and exacerbate MS, as has been pointed out above (Mazzucchelli et al., 2012, Gregory et al., 2007). This hypothesis is in line with the current study's finding (**Figure 4.2**). Whether the IL-7/IL-7Rα pathway agonism/antagonism depends on the IL-7Rα genotype is worth further investigation.

Although my MR results suggest that *IFNGR2, MAPK3, STAT3* and *TYMP* may be of therapeutic benefit for the treatment of MS, the pharmacological effects of the existing medications targeting these genes were not consistent with the direction of their MR effect estimates. Thus, these medications are unlikely suitable for repurposing. Drugs are designed to target genes or pathways in disease-related tissues. Given that gene activities vary across tissues and cell types, a drug's mode of action for the treatment of a specific disease does not necessarily suit another disease, even if these two diseases share the same gene target. This may provide a basis for explaining the inconsistency between the direction of the MR effect estimates of the above genes and the pharmacological effects of the existing medications targeting those genes.

This study found that no drugs have been reported for *GALC* and *MAST3.* This is because these genes were classified as Tier 3 in the druggable genes list (Finan et al., 2017). Of these three genes, *GALC* has been reported to be a risk locus for MS (Consortium*† et al., 2019, Beecham et al., 2013a).

A druggable genome study that integrated Finan's list of druggable genes, MS risk GWAS and QTL data to prioritise druggable targets for MS was recently published (Jacobs et al., 2020). Their approach yielded a list of four prioritised druggable target genes: *CD40*, *MERTK, FCRL3* and *PARP1*. Their final findings did not overlap with the findings of this study. Several methodological differences may explain why these results differ from the current results. Briefly, in the above-mentioned study, the authors employed summary data-based Mendelian randomisation, a technique that was developed to integrate GWAS data and molecular trait data, such as *cis*-eQTL and *cis*-mQTL studies, to determine associations between these traits and outcomes of interest, such as MS (Liu et al., 2021b). The authors prioritise druggable genes in MS if a genetic locus has a causal association between CpG methylation sites and gene expression, association between CpG methylation sites and MS and association between gene expression and MS (Jacobs et al., 2020). On the other hand, my analysis pipeline used the conventional MR approach, only selected SNPs located withing flanking region of 5 kilobases around the encoding gene region and the final list included genes whose expression levels in the blood and/or brain were associated with MS risk.

Similar to MS risk, MR on gene expression measured in blood and brain tissue was employed to identify new drug targets for MS severity. Although several genes reached nominal significance, none passed the significance threshold (FDR $\leq$ 0.05). One explanation could be that the sample size was relatively small; thus, the study did not have the power to pass the significance threshold.

4.4.4. Limitation

One limitation of this work is that MR identifies genes with expression levels associated with the disease. Changes in gene expression may not reflect changes in protein levels and/or account for post-transcriptional and post-translational modifications along the pathway from a gene to a biologically functional protein (Gill et al., 2021). Therefore, genetic variants associated with protein levels (protein quantitative trait loci; pQTL) may model drug target effects more accurately than eQTL (Storm et al., 2021). Due to the lack of pQTL genetic variants for the replicated genes, it was not possible to conduct MR using the pQTL data. Most pQTL studies have been limited by small sample sizes and have mostly involved cell lines rather than primary human tissues (Sun et al., 2018, Yao et al., 2018).

Furthermore, MR in this context might be able to broadly provide information regarding drug class effects, but not necessarily regarding the effects of a specific pharmacological agent (Gill et al., 2021). For example, dihydropyridine and non-dihydropyridine are subclasses of calcium-channel antagonist drugs (i.e., antihypertensive) that exert distinct pharmacological effects. Using human genetic variants within genes corresponding to the targets of antihypertensive drug classes to serve as a proxy for the effects of these treatments can estimate the effects of these drug classes in general but cannot differentiate the relative effects of dihydropyridine versus non-dihydropyridine subclasses(Gill et al., 2019). Here, MR has been used to investigate the effects of perturbing a drug target but is unlikely to be able to offer insight regarding the specific effects of drug subclasses.

MR in this study only captured the on-target effects of the drugs; drug effects that are not exerted through gene targets (off-target effects) cannot be captured in these MR models.

Tissue and cell type-specific traits are crucial to drug discovery and development. This work utilised MR on gene expression measured in blood and brain tissue to capture as many genes as possible and explore two potential tissue sites of action. The majority of identified genes was in blood, and only a few genes were identified in brain tissue, possibly due to the sample size; the blood eQTL data (n = 31,684) was greater than that for brain tissue eQTLs (n = 1,387). Indeed, studies with a larger sample size have greater power to identify many QTL SNPs per gene, which boosts the power of MR results (Huan et al., 2019, Storm et al., 2021).

MS is a heterogeneous disease with many different immune cells contributing to disease pathogenesis (Høglund and Maghazachi, 2014). In this regard, the other important limitation of this work was that it could not distinguish between the effects of changes in the gene expression levels in different types of immune cells on MS due to the QTLs in the blood obtained from the whole blood samples.

It is important to note that these MR findings were based on MS GWAS, which captured the incident risk of the disease (i.e., onset of disease) rather than the progression. Thus, these findings should be interpreted as useful in disease prevention and/or as modulating the effects of the risk of developing MS.

MR estimates reflect the lifelong effects of an exposure, because all exposures instrumented by genetic variants are present from birth and last for a lifetime. Thus, MR does not predict when or at what age the putative risk factors affect the risk of disease nor does it predict the size of the intervention required to reduce the effect of an exposure to prevent disease risk or progression.

4.4.5. Conclusion

In conclusion, this study took advantage of the MR approach and publicly available datasets to provide genetic evidence for *CCR4, SLAMF7* and *SIK3* as possible drug targets for MS.

Table 4.2: Summary of the existing drugs that can target the prioritised genes for MS risk

Gene	Tissue	Drug name	Clinical phase	Indication	Direction of effect for drug	Description	Consistency	Study identifier
CCR4	Blood	Mogamulizumab	Approved	Mycosis fungoides	inhibitor	Block T-cell migration and proliferation	YES	
CD6	Blood	Itolizumab	Approved	Psoriasis	inhibitor	Inhibiting the CD6-CD166 signalling pathway which led to inhibit T cell activation and proliferation	NO	
IFNGR2	Blood	Interferon gamma-1b	Approved	Chronic granulomatous disease	Binder	Initiate a sequence of intracellular events that lead to the transcription of multiple immune-related genes.	NO	
SLAMF7	Blood	Elotuzumab	Approved	Multiple myeloma	modulator	Binding to SLAMF7 on multiple myeloma cells to mediate the killing of these cells	YES	
MAPK3	Blood	Ulixertinib(BVD-523)	Phase 2	Melanoma	inhibitor	Inhibits both ERK 1 and 2, thus preventing the activation of ERK-mediated signal transduction pathways	NO	NCT03417739
STAT3	Blood	OPB-111077	Phase 1	Acute myeloid leukemia/cancer	inhibitor	Not Available	NO	NCT03197714
TYMP	Blood	Tipiracil	Approved	Gastric or colorectal malignancies	Inhibitor	Inhibit TYMP activity	NO	
IL7	Blood and Brain	GSK2618960	Terminated at Phase 1	Multiple sclerosis	Modulator	blocked IL-7 receptor signalling	NO	
IL2RA	Brain	daclizumab	Approved	relapsing forms of multiple sclerosis.	Modulator	blocks the interleukin-2 receptor	YES	NCT01808482
TNFSF14	Blood	Baminercept	phase 2	Rheumatoid arthritis	inhibitor	blockade of LTβ receptor signaling	NO	NCT01552681
SIK3	Blood	Tenalisib	phase 2	Locally Advanced or Metastatic Breast Cancer	Inhibitor	SIK3 inhibitor	YES	NCT05021900

Drug information extracted from Drugbank (*https://www.drugbank.ca*), Therapeutic Target Database (http://db.idrblab.net/ttd/) and ClinicalTrail.gov (https://clinicaltrials.gov/ct2/home). Column labelled with consistency refers to whether the drug direction of effects is consistent with the MR effect estimate. The column labelled with the study identifier refers to the ClinicalTrials.gov identifier. No drug was found for *GALC* and *MAST3*, as these genes were classified as Tiers 3.

Chapter 5 *The role of body fat in multiple sclerosis susceptibility and severity*

This chapter is accepted [March 2022] for publication in ***Multiple Sclerosis Journal***

Statement of contribution

I designed this project, wrote and executed the analysis scripts used in this chapter myself. Nicholas W. Wood, Catherine S. Storm, Demis A. Kia, Rachel Coneys and Burleen K. Chhatwal, contributed to the interpretation of my results.

5.1. Introduction

Obesity is an abnormal or excessive fat accumulation. Obesity is reported as one of the leading causes of death and a risk factor for many metabolic, inflammatory and autoimmune diseases in terms of their incidence, disease severity and outcomes.(Mohammad et al., 2021, Gremese et al., 2014) Therefore, obesity is regarded as a major public health concern responsible for a significant reduction in the quality health (Mohammad et al., 2021). Evidence shows that excess body fat is accompanied by inflammation and alterations in the immune cell function, reflected in an increase in circulating pro-inflammatory proteins, elevated leukocyte, neutrophil, monocyte and lymphocyte counts and impaired immune cell function, leading to an increased risk of severe infectious diseases.(de Heredia et al., 2012, Mohammad et al., 2021)

In obese individuals, research has demonstrated that white adipose tissue (WAT) is the major source of inflammation,(Park et al., 2014) and it can constitute up to 50% of the total body mass in extreme obesity (Kanneganti and Dixit, 2012). WAT stores excess calories in the form of triglycerides and tightly coordinates energy supply and demand at the whole organism level; therefore, WAT is regarded as an energy storage (Kanneganti and Dixit, 2012, Kusminski et al., 2016, Ouchi et al., 2011).

WAT also functions as a key endocrine organ, secreting various adipocyte-derived factors collectively called adipokines, or adipocytokines (Kanneganti and Dixit, 2012, Kusminski et al., 2016, Ouchi et al., 2011). These adipokines can have pro-inflammatory effects, such as interleukin (IL)-1β, IL-6, TNFα and leptin, or anti-inflammatory activities, such as adiponectin (Forny-Germano et al., 2019, Kusminski et al., 2016). Dysregulated secretion of these adipokines, such as in the case of excess adiposity and adipocyte dysfunction, was associated with an increased risk of obese people developing a number of related diseases (Forny-Germano et al., 2019, Ouchi et al., 2011).

During the development of obesity, WAT undergoes expansion due to excessive fat accumulation (hypertrophy) and an increase in the number of adipocytes (hyperplasia) through recruitment of pre-adipocytes from the resident pools of progenitor cells; the latter process embodies *de novo* adipocyte formation (adipogenesis) (Kusminski et al., 2016, Buechler et al., 2015). Due to WAT expansion, vasculature cannot supply oxygen to the expanding tissue, resulting in the activation of hypoxia and hypoxia-inducible factor-1 (HIF-1) (Buechler et al., 2015, Kusminski et al., 2016). Activated HIF-1 inhibits pre-adipocyte differentiation and initiates adipose tissue fibrosis,(Buechler et al., 2015, Kusminski et al., 2016) resulting in limited adipose tissue growth and excess fat stored in ectopic tissues (Buechler et al., 2015). As result of dysfunctional WAT and hypoxia, a shift to an adverse adipokine secretory profile occurs, which typically implies an elevated array of pro-inflammatory factors with a simultaneous reduction in anti-inflammatory factors and recruited monocytes (Kusminski et al., 2016). This pattern of pro-inflammatory cytokine production and immune cell infiltration into adipose tissues establish and maintain a chronic inflammation state (low-grade inflammation). This chronic inflammation is regarded as an aetiology for obesity-related diseases, such as autoimmune diseases, allergy diseases, cancer, atherosclerosis, Alzheimer's disease, insulin resistance and type 2 diabetes (Kuroda and Sakaue, 2017).

Furthermore, ectopic fat deposition in other tissues, such as lymphoid organs (bone marrow and thymus), may occur because WAT becomes severely dysfunctional and unable to expand properly to store surplus energy (Kusminski et al., 2016, Mohammad et al., 2021). Increased fat deposition in these tissues leads to changes in the distribution of leukocyte populations, lymphocyte activity and overall immune defences, which may weaken the immune system and make obese patients vulnerable to infectious diseases (Andersen et al., 2016, Mohammad et al., 2021).

These data, indeed, raise a critical question of whether losing weight can reduce inflammation and therefore the risk of obesity-related diseases? Interestingly, evidence revealed that weight-loss interventions are a determining factor for reducing the number of adipose tissue macrophages, the level of pro-inflammatory profiles in obese individuals and incident diseases risk (Ouchi et al., 2011, Forsythe et al., 2008). For example, a systematic review showed that a 1-kg loss of body weight was associated with a −0·13 mg/l change in C-reactive protein, achieved through diet and lifestyle modifications, which increased to a -0·16 mg/l change per 1 kg weight loss caused by gastric surgery (Forsythe et al., 2008, Selvin et al., 2007). Furthermore, studies related to the effects of weight loss on disease risk and activity have shown promising results. For example Klingberg et al. showed that weight loss treatment with a very low energy diet significantly reduced the disease activity in joints, entheses and skin in patients with psoriatic arthritis and obesity (Klingberg et al., 2019). Additionally, Haase et al. found in their study that 13% loss of body weight in obese patients is associated with lower risk of type 2 diabetes, chronic kidney disease, hypertension and dyslipidaemia (Haase et al., 2021).

5.2. The role of body fat in multiple sclerosis

In MS, a link has been demonstrated between obesity, the risk of developing MS and a worsening disability level in MS patients. Recent observational and MR studies have shown that obesity in childhood or early adulthood, as measured by elevated BMI, is associated with the risk of developing MS (Munger et al., 2009, Hedström et al., 2012, Mokry et al., 2016, Harroud et al., 2021). Although BMI is commonly used to identify obese persons due to its ease of calculation and cost-effectiveness (Pilutti and Motl, 2019), its use has been criticised (Frankenfield et al., 2001). BMI does not distinguish between the contributions of fat and non-fat tissue (e.g., lean tissue mass) to body weight (Speed et al., 2019) which might contribute to the misclassification of certain groups of individuals. For instance, athletes might be classified as obese due to their higher BMI, but the BMI increase comes from higher lean muscle mass, not from accumulated fat (Jonnalagadda et al., 2004).

In addition, BMI does not capture the location of body fat and non-fat, which have been shown to play an essential role in predicting the risk of several diseases. For example, increased abdominal fat is associated with cancer, stroke and cardiometabolic disease, whereas decreased lean mass of the arms and legs is associated with increased falls and frailty (Wingo et al., 2018).

Although MR studies have been employed to investigate the relation between obesity (BMI) and MS risk, the MR approach has not yet been used to investigate the relation between obesity and the progression of disability in MS. Observational studies of obesity and the progression of disability in MS patients have reported inconsistent findings. Whereas some studies report evidence supporting the association between higher BMI,FM, FP and disability progression in MS (Pilutti and Motl, 2019, Richter et al., 2017), other studies have identified no evidence of an association between disability progression in MS and BMI, FM or FP (Lambert et al., 2002, Tadić et al., 2020).

5.3. Aims

I aimed to conduct an MR analysis to investigate the relationship between MS and anthropometric measures. Specifically, I assessed two questions:

1. The causal effects of lifelong genetically elevated adult BMI, height, weight, FM, FP and NFM — henceforth, “anthropometric measures” — on MS risk and severity to better understand the effect of obesity on MS.
2. Whether the lifelong genetic predisposition to increased MS risk affects anthropometric measures.

5.4. Method

To test whether each anthropometric measure is a causal risk factor for MS risk and severity, I obtained the summary statistics of 21 anthropometric measures from the Neale Lab consortium. These 21 anthropometric measures were divided into two categories (**Table 5.1**). The adiposity-related measures included BMI, weight, FP and FM for the whole body, the upper limbs (the right and the left arms), the lower limbs (the right and the left legs) and the trunk. The second category included the height and the NFM for the whole body, upper limbs, lower limbs and trunk. With anthropometric measure-associated SNPs regarded as exposure (method section 2.2.2.3), I then obtained the corresponding effect estimates for MS risk and MS severity as the outcome from IMSGC (method section 2.2.3).

To explore whether MS influences anthropometric measures (bidirectional MR analysis), I selected 200 susceptibility variants as GWAS for MS risk obtained from the discovery phase of the most recent cohort of the IMSGC (method section 2.2.3) and then obtained the corresponding effect estimates for anthropometric measures as the outcome from the Neale Lab consortium.

All the selected SNPs were clumped at LD = $r^2 < 0.001$, followed by Steiger filtering. Then, the mean F-statistic and R^2 were computed. For the statistical analysis, I used IVW as the main estimator and MR-Egger, weighted median, Cochran's Q and I^2 tests and the MR-Egger intercept as sensitivity tools. I also used IVW MR-radial to identify the outlier SNPs. Finally, MVMR was used to account for pleiotropy and estimate the direct effect. For more details about the selection and validity assessment of the genetic variants and the statistical method, refer to the method chapter.

5.4. Results

Table 5.1 shows the exact numbers of genetic variants, sample size, R^2, the mean F-statistic and the means SD for the anthropometric measures. Weak instrument bias is likely negligible in this data since the mean F-statistic is greater than 10. Figure 5.1 shows the genetic correlations among the 21 anthropometric measures.

Table 5.1. Sample characteristics for the traits of interest

Exposure (UK biobank code)	Mean (standard deviation)	sample size	Anthropometrics-MS risk			Anthropometrics-MS severity		
			Independent genome-wide significant SNPs	Approximate variance explained (%)	mean *F*-statistic	Independent genome-wide significant SNPs	Approximate variance explained (%)	mean *F*-statistic
MS risk (as exposure)		115,803	97	19[a]	77			
Height (50)	168.5 cm (9.3)	360,388	530	8.6	118	401	6.9	126
BMI (21,001)	27.3 kg/m2 (4.8)	359,983	266	4.2	58	168	2.9	64
Weight (21,002)	77.9 kg (15.9)	360,116	304	4.2	64	168	2.9	64
Trunk fat-mass (23,128)	13.7 kg (5.1)	354,597	261	4.1	56	176	3.1	62
Arm fat mass left (23,124)	1.3 kg (0.7)	354,673	249	3.8	57	165	2.7	62
Arm fat mass right (23,120)	1.2 kg (0.6)	354,736	253	3.9	57	162	2.7	63
Arm fat percentage left (23,123)	30.2 % (10.2)	354,707	225	2	55	151	1.4	60
Arm fat percentage right (23,119)	29.4 % (10.1)	354,760	236	2.1	55	147	1.4	60
Arm non-fat mass left (23,125)	2.9 kg (0.8)	354,668	330	2.6	69	228	1.9	75
Arm non-fat mass right (23,121)	2.9 kg (0.8)	354,732	323	2.5	71	218	1.8	77
Whole body fat mass (23,100)	24.8 kg (9.5)	354,244	258	3.8	56	168	2.8	63
Whole body fat percentage (23,099)	31.4 % (8.5)	354,628	240	2.1	54	156	1.4	58
Whole body non-fat mass (23,101)	53.1 kg (11.5)	354,808	374	3.1	76	242	2.2	84
Leg fat mass left (23,116)	4.2 kg (1.9)	354,788	249	2.5	56	155	1.7	63
Leg fat mass right (23,112)	4.3 kg (1.9)	354,807	249	2.5	57	152	1.7	64
Leg fat percentage left (23,115)	32.0 % (10.6)	354,791	228	1.2	52	144	0.8	54
Leg fat percentage right (23,111)	32.0 % (10.7)	354,811	230	1.3	52	149	0.9	54
Leg non-fat mass left (23,117)	8.8 kg (2.0)	354,771	333	2.8	72	217	1.9	78
Leg non-fat mass right (23,113)	8.9 kg (2.0)	354,798	350	2.9	71	228	2	77
Trunk fat percentage (23,127)	31.1 % (8.0)	354,619	225	2.8	54	148	2	59
Trunk non-fat mass (23,129)	29.6 kg (6.0)	354,530	353	3	79	243	2.3	86

[a] Because MS risk is a binary trait, variance explained was extracted from the original MS risk genome-wide association study.

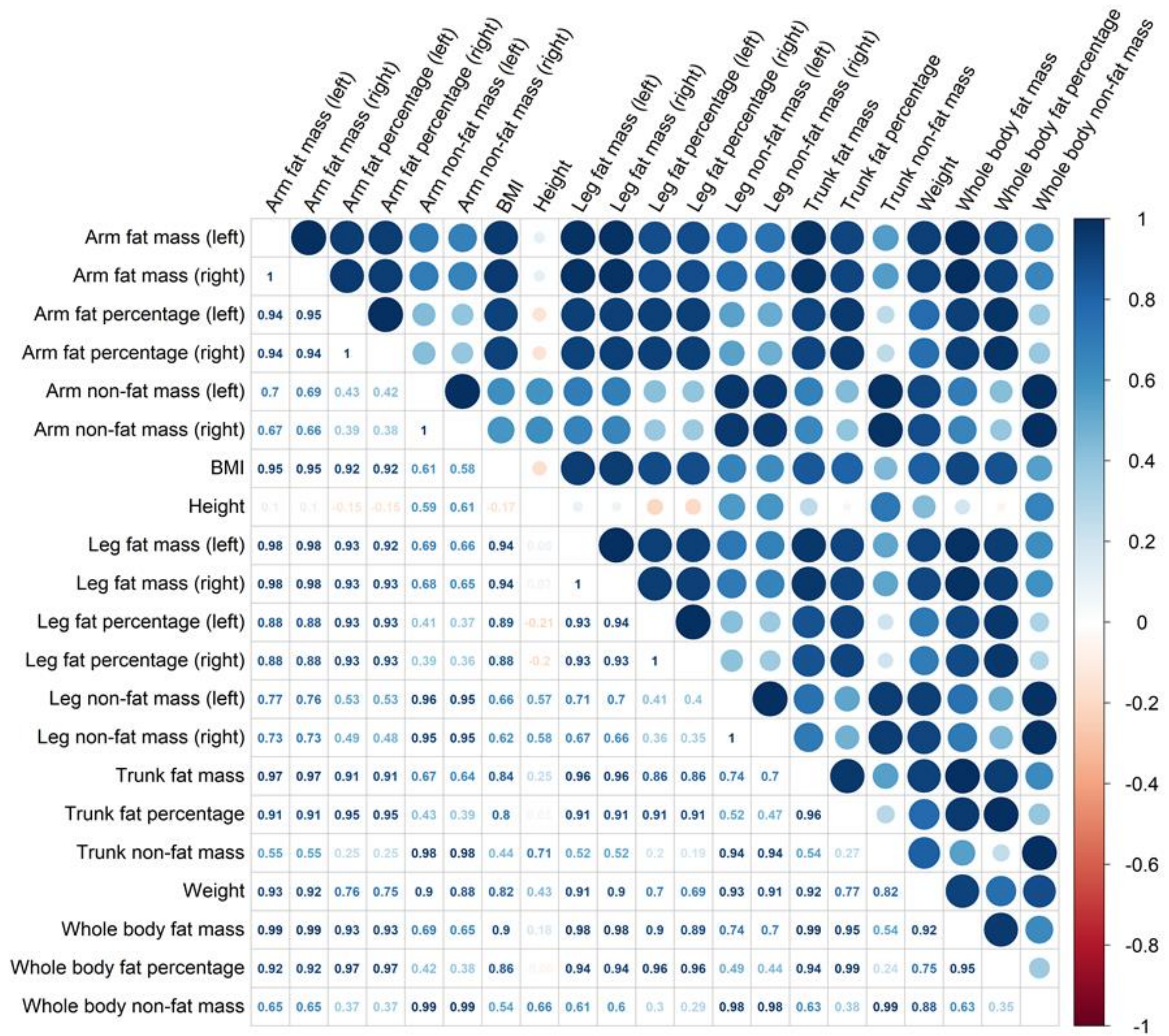


Figure 5.1: Genetic correlations between the 21 anthropometric measures. The heat scale represents the strength of genetic correlation, with blue colour indicating positive genetic correlation and red colour indicating negative genetic correlation. Light colours correspond to lower correlations, while darker colours correspond to stronger correlations.

5.4.1. Influence of genetically raised anthropometric measures on the risk of MS

Table 5.2 displays the odds ratio for MS risk per one SD increase in each of the anthropometric measures. The IVW results showed that genetically raised BMI, weight, FM and FP in the whole body, trunk, arms (left and right) and legs (left and right) were causally associated with an increased MS risk. For the sensitivity analyses, the results from the weighted median method further replicated the direction and significance of the IVW results, thereby providing additional confidence in the IVW results. The results from MR-Egger were also similar to the IVW results in the direction of the causal associations with a moderate increase in the effect estimates; however, CI were wider, resulting in a number of estimations crossing the null. The I^2 statistic indicated a slight degree of heterogeneity; however, Cochran's Q *p* values were not significant, except for the weight. The MR-Egger intercept indicated no evidence for horizontal pleiotropy except for the FM in the right and left legs where the intercept *p* values were significant (p-value < 0.05), suggests horizontal pleiotropy effect. The pleiotropy-corrected causal estimates from the MR-Egger and weighted median for these measures were moderately increased relative to the IVW estimates, but still significant and in the same direction, further supporting the causal role of FM in the legs on MS risk. For height and NFM in the whole body, trunk, arms and legs, the MR results found no evidence for a relationship between these measures and MS risk.

For the MVMR analyses, we fitted a model with adiposity-related measures that retained an effect on MS risk in the univariable MR models. The MVMR-IVW revealed that compared with univariable estimates, the direct estimates were slightly lower for the BMI, weight, FM and FP at different body parts and slightly larger for the FP in the right arm and the FM in the trunk but still significant. Meanwhile, the direct estimates for the FM in the left leg were attenuated, resulting in a wider 95% CI that overlapped null. Thus, the observed effects for left leg FM in the univariable MR analyses are more likely operating through the pathways of other adiposity-related measures

Table 5.2: MR estimates for the effect of anthropometric measures on MS risk (significance level is set at IVW FDR ≤ 0.05)

Exposure	No. of SNPs	IVW			MR-Egger		Weighted median		pleiotropy assessment					MVMR-IVW		
		OR (95%CI)	pval	FDR	OR (95%CI)	pval	OR (95%CI)	pval	Q	Q pval	I^2(%)	MR-Egger intercept	MR-Egger intercept pval	No. of SNPs	OR (95%CI)	pval
Whole body fat percentage	240	1.49 (1.25,1.78)	**6.36E-06**	**1.34E-04**	1.88 (1.01,3.5)	**4.86E-02**	1.6 (1.24,2.07)	**3.27E-04**	271.109	6.91E-02	12.2	-0.0035	0.453	224	1.48 (1.26,1.75)	**2.60E-06**
Leg fat percentage (left)	228	1.52 (1.21,1.9)	**2.74E-04**	**1.17E-03**	2.65 (1.15,6.14)	**2.35E-02**	1.64 (1.16,2.33)	**5.25E-03**	257.13	7.59E-02	12.1	-0.0068	0.176	217	1.41 (1.13,1.77)	**2.30E-03**
Leg fat percentage (right)	230	1.48 (1.19,1.84)	**4.58E-04**	**1.38E-03**	2.57 (1.15,5.74)	**2.21E-02**	1.6 (1.17,2.19)	**3.64E-03**	258.177	8.29E-02	11.7	-0.0068	0.163	215	1.33 (1.08,1.64)	**7.90E-03**
Arm fat percentage (right)	236	1.39 (1.18,1.64)	**1.02E-04**	**1.07E-03**	1.67 (0.98,2.84)	6.01E-02	1.58 (1.22,2.06)	**5.83E-04**	245.817	2.85E-01	4.8	-0.0029	0.477	223	1.49 (1.27,1.76)	**1.60E-06**
Arm fat percentage (left)	225	1.3 (1.1,1.55)	**2.57E-03**	**5.39E-03**	1.96 (1.14,3.37)	**1.50E-02**	1.51 (1.15,1.98)	**2.75E-03**	239.56	2.13E-01	6.9	-0.0064	0.119	212	1.23 (1.04,1.45)	**1.46E-02**
Trunk fat percentage	225	1.29 (1.11,1.49)	**7.24E-04**	**1.90E-03**	1.46 (0.87,2.45)	1.53E-01	1.42 (1.14,1.76)	**1.67E-03**	243.956	1.60E-01	8.6	-0.0023	0.622	211	1.26 (1.1,1.45)	**1.22E-03**
Whole body fat mass	258	1.23 (1.09,1.4)	**1.15E-03**	**2.68E-03**	1.35 (0.91,2)	1.32E-01	1.28 (1.06,1.55)	**1.00E-02**	289.278	7.49E-02	11.5	-0.0019	0.627	244	1.21 (1.07,1.37)	**1.97E-03**
Arm fat mass (left)	249	1.26 (1.11,1.42)	**2.78E-04**	**1.17E-03**	1.3 (0.9,1.86)	1.61E-01	1.3 (1.06,1.59)	**1.29E-02**	258.662	2.92E-01	4.5	-0.0006	0.857	237	1.24 (1.1,1.4)	**6.24E-04**
Arm fat mass (right)	253	1.24 (1.1,1.4)	**4.40E-04**	**1.38E-03**	1.4 (0.98,1.99)	6.80E-02	1.29 (1.06,1.56)	**1.12E-02**	260.027	3.34E-01	3.5	-0.0024	0.497	243	1.2 (1.07,1.36)	**2.83E-03**
Leg fat mass (right)	249	1.22 (1.05,1.42)	**9.64E-03**	**1.69E-02**	2.14 (1.35,3.39)	**1.36E-03**	1.33 (1.04,1.7)	**2.51E-02**	255.956	3.34E-01	3.5	-0.0094	**0.0121**	237	1.19 (1.03,1.38)	**2.18E-02**
Leg fat mass (left)	249	1.2 (1.03,1.39)	**2.08E-02**	**3.11E-02**	2.01 (1.28,3.17)	**2.88E-03**	1.28 (1.01,1.64)	**4.22E-02**	250.115	4.33E-01	1.2	-0.0086	**0.0184**	238	1.15 (0.99,1.34)	6.37E-02
Trunk fat mass	261	1.16 (1.03,1.31)	**1.52E-02**	**2.46E-02**	1.28 (0.86,1.89)	2.20E-01	1.26 (1.05,1.52)	**1.42E-02**	285.926	1.20E-01	9.4	-0.002	0.621	247	1.18 (1.05,1.33)	**5.48E-03**
BMI	266	1.25 (1.11,1.4)	**1.89E-04**	**1.17E-03**	1.6 (1.15,2.21)	**5.25E-03**	1.23 (1.01,1.5)	**4.45E-02**	275.14	3.06E-01	4	-0.0053	0.114	253	1.21 (1.08,1.36)	**8.85E-04**
Weight	304	1.2 (1.05,1.36)	**5.92E-03**	**1.13E-02**	1.3 (0.91,1.84)	1.47E-01	1.21 (1,1.46)	**4.80E-02**	365.459	7.20E-03	17.4	-0.0016	0.634	281	1.13 (1.01,1.28	**3.50E-02**
Height	530	1.02 (0.92,1.12)	7.40E-01	7.40E-01	1.11 (0.9,1.38)	3.41E-01	1.02 (0.88,1.17)	8.25E-01	715.734	8.63E-08	26.2	-0.002	0.369			
Whole body non-fat mass	374	1.11 (0.95,1.31)	1.82E-01	2.12E-01	1.27 (0.83,1.93)	2.66E-01	1.09 (0.87,1.36)	4.57E-01	510.215	2.42E-06	27.1	-0.002	0.513			
Leg non-fat mass (left)	333	1.15 (0.98,1.37)	9.44E-02	1.32E-01	1.06 (0.68,1.67)	7.87E-01	1.16 (0.91,1.48)	2.28E-01	467.398	1.10E-06	29.2	0.0013	0.703			
Leg non-fat mass (right)	350	1.15 (0.97,1.35)	1.04E-01	1.37E-01	1.1 (0.71,1.71)	6.70E-01	1.1 (0.87,1.4)	4.36E-01	475.552	6.22E-06	26.8	0.0006	0.841			
Arm non-fat mass (right)	323	1.09 (0.91,1.31)	3.50E-01	3.87E-01	1.18 (0.71,1.97)	5.18E-01	1.1 (0.85,1.42)	4.62E-01	454.561	1.28E-06	29.4	-0.0012	0.737			
Arm non-fat mass (left)	330	1.08 (0.91,1.29)	3.79E-01	3.98E-01	1.18 (0.72,1.93)	5.18E-01	1.1 (0.85,1.42)	4.63E-01	460.522	1.82E-06	28.8	-0.0013	0.722			
Trunk non-fat mass	353	1.13 (0.96,1.32)	1.46E-01	1.81E-01	0.99 (0.65,1.51)	9.57E-01	1.09 (0.86,1.39)	4.89E-01	476.228	9.22E-06	26.3	0.0021	0.515			

IVW = inverse variance-weighted (multiplicative-random effect); No. of SNPs = number of single nucleotide polymorphisms that included in the analysis; OR = odds ratio; 95% CI = confidence interval; pval = nominal p-value; FDR = False discovery rate; Q = cochran's Q test; Q pval = Cochran's Q test p-value; I^2 (%) = expresses the level of heterogeneity as a **percentage**; MR-Egger Intercept pval = MR-Egger Intercept p value; MVMR= multivariable mendelian randomization.

5.4.2. Influence of genetically raised MS risk on anthropometric measures

We further conducted a bidirectional MR to assess the causal relationship between MS risk and anthropometric measures, as shown in Table 5.3, which displays the β-coefficients for each anthropometric measure per log odds increase in MS risk. After removing six genetic instruments not found in the anthropometric data, 37 genetic instruments with incompatible alleles and one genetic instrument for being palindromic, a total of 97 instruments were used for this analysis. The IVW and MR-Egger results revealed that a genetic predisposition to MS has no significant effect on any of the anthropometric measures investigated here, except for the weighted median, where the p-values for some of the anthropometric measures were significant. Since there was no evidence of pleiotropy from the MR-Egger intercept, IVW is more robust for detecting the true causal effect than the weighted median. There was significant evidence of heterogeneity as reflected by the I^2 statistic and Cochran's Q p-values. Since the pleiotropy is balanced, the heterogeneity is more likely due to the non-collapsibility of the odds ratio (Hemani et al., 2018a). Further, the heterogeneity in these analyses was accounted for by using the IVW multiplicative random-effects model (Burgess et al., 2019)

Table 5.3: MR estimates for the effect of MS risk on anthropometric measures (bidirectional analysis)

Outcome	No. of SNPs	IVW			MR-Egger		Weighted median		pleiotropy assessment				
		β (95%CI)	pval	FDR	β (95%CI)	pval	β (95%CI)	pval	Q	Q pval	I^2 (%)	MR-Egger intercept	MR-Egger Intercept pval
Whole body fat percentage	97	-0.0027(-0.0106,-0.0106)	4.99E-01	5.82E-01	0.0013(-0.0219,-0.0219)	9.15E-01	-0.0045(-0.0116,-0.0116)	2.13E-01	362.819	6.23E-33	73.8	-0.0004	0.721
Arm fat percentage (left)	97	-0.0052(-0.013,-0.013)	1.91E-01	3.67E-01	7e-04(-0.0221,-0.0221)	9.52E-01	-0.0126(-0.0197,-0.0197)	**4.84E-04**	350.816	5.33E-31	72.9	-0.0006	0.593
Arm fat percentage (right)	97	-0.0043(-0.0121,-0.0121)	2.80E-01	4.20E-01	0.0037(-0.0192,-0.0192)	7.52E-01	-0.0107(-0.0178,-0.0178)	**2.98E-03**	351.458	4.21E-31	73	-0.0009	0.468
Leg fat percentage (left)	97	-0.0021(-0.0087,-0.0087)	5.26E-01	5.82E-01	-0.0026(-0.0221,-0.0221)	7.93E-01	-0.0053(-0.0111,-0.0111)	7.05E-02	386.698	7.72E-37	75.4	0.0001	0.96
Leg fat percentage (right)	97	-0.0016(-0.0083,-0.0083)	6.38E-01	6.69E-01	-0.0025(-0.0221,-0.0221)	8.03E-01	-0.0041(-0.0097,-0.0097)	1.54E-01	382.452	3.87E-36	75.2	0.0001	0.924
Trunk fat percentage	97	-0.003(-0.0122,-0.0122)	5.26E-01	5.82E-01	0.0055(-0.0216,-0.0216)	6.91E-01	-0.0077(-0.0163,-0.0163)	8.15E-02	342.521	1.12E-29	72.3	-0.0009	0.515
Whole body fat mass	97	-0.0061(-0.0156,-0.0156)	2.10E-01	3.67E-01	0.0049(-0.023,-0.023)	7.31E-01	-0.0075(-0.0161,-0.0161)	8.52E-02	315.268	2.02E-25	69.9	-0.0012	0.414
Arm fat mass (left)	97	-0.0078(-0.0171,-0.0171)	9.99E-02	3.67E-01	0.0042(-0.023,-0.023)	7.63E-01	-0.0111(-0.0203,-0.0203)	1.77E-02	293.994	3.37E-22	67.7	-0.0013	0.361
Arm fat mass (right)	97	-0.0071(-0.0163,-0.0163)	1.34E-01	3.67E-01	0.0063(-0.0208,-0.0208)	6.47E-01	-0.0068(-0.0156,-0.0156)	1.30E-01	291.345	8.35E-22	67.4	-0.0014	0.304
Leg fat mass (left)	97	-0.0041(-0.012,-0.012)	3.05E-01	4.26E-01	0.0012(-0.022,-0.022)	9.22E-01	-0.0074(-0.0146,-0.0146)	**4.35E-02**	330.645	8.32E-28	71.3	-0.0006	0.635
Leg fat mass (right)	97	-0.0036(-0.0115,-0.0115)	3.67E-01	4.82E-01	0.0012(-0.022,-0.022)	9.20E-01	-0.0075(-0.0146,-0.0146)	**3.94E-02**	322.935	1.32E-26	70.6	-0.0005	0.666
Trunk fat mass	97	-0.0059(-0.0157,-0.0157)	2.40E-01	3.87E-01	0.0063(-0.0225,-0.0225)	6.67E-01	-0.0032(-0.0121,-0.0121)	4.77E-01	317.838	8.13E-26	70.1	-0.0013	0.378
BMI	97	-0.0072(-0.0171,-0.0171)	1.57E-01	3.67E-01	0.0045(-0.0246,-0.0246)	7.63E-01	-0.0086(-0.0177,-0.0177)	6.33E-02	333.353	3.13E-28	71.5	-0.0013	0.406
Weight	97	-0.0074(-0.0168,-0.0168)	1.27E-01	3.67E-01	0.0083(-0.0193,-0.0193)	5.56E-01	-0.0032(-0.0112,-0.0112)	4.24E-01	381.51	5.54E-36	75.1	-0.0017	0.239
Height	97	-0.0021(-0.0155,-0.0155)	7.59E-01	7.59E-01	0.0089(-0.0307,-0.0307)	6.62E-01	-2e-04(-0.0068,-0.0068)	9.58E-01	1225.761	7.96E-196	92.2	-0.0012	0.565
Whole body non-fat mass	97	-0.0066(-0.0152,-0.0152)	1.34E-01	3.67E-01	0.0045(-0.0208,-0.0208)	7.28E-01	-0.0069(-0.0127,-0.0127)	**1.91E-02**	630.768	5.25E-80	84.9	-0.0012	0.363
Arm non-fat mass (left)	97	-0.0055(-0.0138,-0.0138)	1.97E-01	3.67E-01	0.0038(-0.0206,-0.0206)	7.60E-01	-0.004(-0.0099,-0.0099)	1.90E-01	571.337	4.31E-69	83.4	-0.001	0.429
Arm non-fat mass (right)	97	-0.0064(-0.0146,-0.0146)	1.27E-01	3.67E-01	0.0014(-0.0226,-0.0226)	9.08E-01	-0.0083(-0.0139,-0.0139)	**3.43E-03**	577.485	3.27E-70	83.5	-0.0008	0.501
Leg non-fat mass (left)	97	-0.0061(-0.0138,-0.0138)	1.20E-01	3.67E-01	0.0078(-0.0148,-0.0148)	4.98E-01	-0.0026(-0.0083,-0.0083)	3.69E-01	473.153	1.46E-51	79.9	-0.0015	0.2
Leg non-fat mass (right)	97	-0.0069(-0.0147,-0.0147)	8.56E-02	3.67E-01	0.0067(-0.0162,-0.0162)	5.69E-01	-0.005(-0.0109,-0.0109)	9.47E-02	491.258	9.72E-55	80.7	-0.0015	0.22
Trunk non-fat mass	97	-0.0067(-0.0159,-0.0159)	1.55E-01	3.67E-01	0.0035(-0.0234,-0.0234)	7.97E-01	-0.0067(-0.0124,-0.0124)	**2.25E-02**	712.497	2.66E-95	86.7	-0.0011	0.431

For Abbreviations, see Table 5.2.

5.4.3. Influence of genetically raised anthropometric measures on MS severity

Table 5.4 displays the β-coefficients for MS severity per one SD increase in each anthropometric measure. The IVW results showed that genetically raised BMI, weight and FM in the whole body, trunk, arms and legs were causally associated with an increase in MS severity. For sensitivity analyses, the estimates were slightly increased in the MR-Egger estimator, while they were nearly identical to the IVW in the weighted median estimator. The MR-Egger and weighted median replicated the IVW direction of the estimates but did not reach statistical significance due to wide CIs. The Cochran's Q, I^2 statistic and MR-Egger intercept indicated no evidence for heterogeneity or horizontal pleiotropy. Thus, causal estimates were more convincing in the IVW results.

For the FP, the MR findings revealed evidence that genetically raised trunk FP is causally associated with an increase in MS severity; however, it did not pass the FDR. By contrast, we did not detect any significant association between FP in the whole body or the other limbs (arms and legs) and MS severity. The power to detect a significant association here would seem to be low due to the small proportion of variance explained by the genetic variants associated with FP in the whole body and limbs ($R^2 = 0.8–1.4\%$) compared with the corresponding values for other anthropometric measures that retained an effect on MS severity (**Table 5.1**). For height and NFM in the whole body, trunk, arms and legs, we found no evidence of the causal role of these measures on MS severity.

We took the measures that retained an effect on MS severity in the univariable MR model forward and further fitted an MVMR model. The MVMR-IVW revealed that BMI, weight, and FM in the whole body, trunk, legs and left arm have a significant direct effect on MS severity, but the estimates are slightly lower than the total estimates in the univariable MR analyses. The direct effect for FM in the right arm attenuated to the null after adjusting for the other adiposity-related measures suggests that the effects of FM in the right arm on MS severity were more likely operating through the pathways of other adiposity-related measures.

Table 5.4: MR estimates for the effect of anthropometric measures on MS severity (significance level is set at IVW FDR ≤ 0.05)

Exposure	No. of SNPs	IVW			MR-Egger		Weighted median		pleiotropy assessment					MVMR-IVW		
		β (95%CI)	pval	FDR	β (95%CI)	pval	β (95%CI)	pval	Q	Q pval	I^2(%)	MR-Egger intercept	MR-Egger intercept pval	No. of SNPs	OR (95%CI)	pval
Whole body fat mass	168	0.35 (0.13,0.57)	**1.74E-03**	**6.09E-03**	0.8 (-0.34,1.94)	1.72E-01	0.4 (-0.22,1.02)	2.03E-01	51.305	1	0	-0.0095	0.414	141	0.25 (0.01,0.5)	**0.04**
Leg fat mass (right)	152	0.48 (0.18,0.77)	**1.55E-03**	**6.09E-03**	0.99 (-0.44,2.43)	1.76E-01	0.43 (-0.38,1.24)	2.95E-01	51.714	1	0	-0.0092	0.451	124	0.28 (-0.004,0.57)	**0.05**
Leg fat mass (left)	155	0.47 (0.17,0.76)	**2.10E-03**	**6.31E-03**	0.89 (-0.55,2.33)	2.28E-01	0.47 (-0.35,1.3)	2.60E-01	53.21	1	0	-0.0074	0.539	147	0.32 (0.02,0.62)	**0.04**
Arm fat mass (left)	165	0.37 (0.15,0.59)	**7.86E-04**	**6.09E-03**	0.74 (-0.36,1.83)	1.89E-01	0.38 (-0.26,1.02)	2.43E-01	48.231	1	0	-0.0079	0.483	165	0.26 (0.04,0.48)	**0.02**
Arm fat mass (right)	162	0.3 (0.08,0.52)	**7.18E-03**	**1.89E-02**	0.8 (-0.31,1.92)	1.58E-01	0.35 (-0.28,0.99)	2.73E-01	47.483	1	0	-0.0109	0.342	162	0.2 (-0.03,0.42)	0.08
Trunk fat mass	176	0.35 (0.14,0.56)	**1.24E-03**	**6.09E-03**	0.73 (-0.37,1.84)	1.95E-01	0.42 (-0.16,1)	1.55E-01	55.081	1	0	-0.0084	0.469	93	0.31 (0.06,0.56)	**0.01**
BMI	168	0.36 (0.14,0.57)	**1.04E-03**	**6.09E-03**	0.92 (-0.11,1.96)	8.23E-02	0.34 (-0.3,0.97)	2.95E-01	50.741	1	0	-0.0127	0.25	168	0.28 (0.06,0.51)	**0.01**
Weight	168	0.36 (0.14,0.57)	**1.04E-03**	**6.09E-03**	0.92 (-0.11,1.96)	8.23E-02	0.34 (-0.3,0.98)	3.01E-01	50.741	1	0	-0.0127	0.25	127	0.26 (0.02,0.49)	**0.03**
Whole body fat percentage	156	0.18 (-0.11,0.48)	2.24E-01	4.71E-01	0.49 (-1.41,2.38)	6.16E-01	0.42 (-0.35,1.19)	2.88E-01	45.049	1	0	-0.0047	0.743			
Arm fat percentage (left)	151	0.15 (-0.15,0.46)	3.33E-01	5.83E-01	1.04 (-0.68,2.76)	2.39E-01	0.26 (-0.59,1.12)	5.47E-01	45.712	1	0	-0.0141	0.288			
Arm fat percentage (right)	147	0.08 (-0.22,0.38)	6.20E-01	7.73E-01	0.92 (-0.84,2.69)	3.08E-01	0.13 (-0.72,0.98)	7.65E-01	41.891	1	0	-0.0135	0.325			
Leg fat percentage (right)	149	0.1 (-0.26,0.46)	5.79E-01	7.73E-01	0.58 (-1.98,3.14)	6.56E-01	0.25 (-0.77,1.28)	6.29E-01	37.58	1	0	-0.006	0.702			
Leg fat percentage (left)	144	0.02 (-0.34,0.39)	8.98E-01	9.03E-01	0.28 (-2.29,2.85)	8.29E-01	0.21 (-0.85,1.27)	7.02E-01	35.909	1	0	-0.0032	0.836			
Trunk fat percentage	148	0.27 (0.01,0.52)	**4.36E-02**	1.02E-01	0.22 (-1.37,1.82)	7.83E-01	0.52 (-0.14,1.17)	1.21E-01	45.232	1	0	0.0008	0.958			
Whole body non-fat mass	242	-0.11 (-0.39,0.17)	4.44E-01	6.66E-01	0.09 (-1.07,1.25)	8.78E-01	-0.37 (-1,0.27)	2.55E-01	94.515	1	0	-0.0033	0.717			
Arm non-fat mass (right)	218	-0.07 (-0.37,0.22)	6.26E-01	7.73E-01	0.3 (-1.06,1.65)	6.68E-01	-0.26 (-0.95,0.43)	4.62E-01	79.801	1	0	-0.0057	0.565			
Arm non-fat mass (left)	228	-0.02 (-0.31,0.28)	9.03E-01	9.03E-01	0.25 (-1.06,1.56)	7.10E-01	-0.1 (-0.81,0.62)	7.91E-01	87.229	1	0	-0.0042	0.668			
Leg non-fat mass (left)	217	0.06 (-0.24,0.36)	6.90E-01	8.05E-01	0.59 (-0.66,1.85)	3.54E-01	-0.05 (-0.74,0.63)	8.77E-01	85.704	1	0	-0.0087	0.369			
Leg non-fat mass (right)	227	0.02 (-0.27,0.31)	8.87E-01	9.03E-01	0.28 (-0.96,1.52)	6.59E-01	-0.12 (-0.82,0.57)	7.28E-01	90.03	1	0	-0.0042	0.661			
Trunk non-fat mass	243	-0.13 (-0.41,0.15)	3.61E-01	5.83E-01	-0.13 (-1.29,1.03)	8.26E-01	-0.36 (-1.01,0.29)	2.80E-01	100.419	1	0	0	0.999			
Height	401	-0.08 (-0.26,0.09)	3.54E-01	5.83E-01	-0.39 (-0.98,0.21)	2.04E-01	-0.11 (-0.5,0.27)	5.63E-01	196.563	1	0	0.0069	0.272			

For abbreviation see Table 5.4

5.5. Discussion

The purpose of the present study was to explore the causal role of anthropometric measures on MS to obtain a better understanding of the impact of excessive fat (obesity) and non-fat mass on MS risk and severity. Our study provides evidence from human genetics that obesity-related measures are an important contributor to MS development and greater disability progression, but height and NFM are not.

Our MR findings first confirmed that a higher BMI leads to a greater risk of developing MS but found no evidence that MS risk influences BMI or the other anthropometric measures. This finding supports previous observational and MR studies that found an association between elevated BMI and an increase in the risk of developing MS (Mokry et al., 2016, Hedström et al., 2012, Munger et al., 2009, Harroud et al., 2021).

BMI does not differentiate between fat and non-fat tissues and between fat stored in different parts of the body. Thus, BMI can partially be used to study obesity. Therefore, we used a range of anthropometric measures that enable capture of the fat stored in different compartments of the body and to discriminate between fat and non-fat tissues. Our MR findings suggest that people with greater fat stored in the whole body, arms, legs and trunk are at a high risk of developing MS. On the other hand, our findings indicated that height or having an increase in NFM are unlikely to put an individual at high risk of getting MS. This lack of evidence of associations between height, NFM and MS risk is unlikely due to low power since the number of GWAS-associated genetic instruments for height and NFM and the variance they explained are greater than the corresponding values—for example, for FP or FM in the legs, which retained a significant causal association with MS risk.

Of course, recommending a healthy weight (or BMI) to all people is important for lowering their risk of a host of diseases, and we can now add MS to this list. However, in terms of advising and managing patients who already have a diagnosis of MS, understanding whether fat and/or non-fat mass may play a role is perhaps even more important. Therefore, we were particularly interested in identifying any causal effect of anthropometric measures on MS severity. Our MR findings suggest that obesity is a significant contributor to disability progression, and therefore severity, in MS, as evident by MR results for BMI, weight and FM. These findings support previous observational studies that found that higher BMI and fat accumulation in the whole body, arm, leg and trunk are significantly associated with greater disability in MS patients (Pilutti and Motl, 2019, Richter et al., 2017). Our findings are also in line with observational studies that identified no link between NFM and disability progression in MS (Pilutti and Motl, 2019). By contrast, our findings are in disagreement with the results of other observational studies that identified no evidence of an association between disability progression in MS and adiposity-related measures, with respect to BMI, FM or FP (Lambert et al., 2002, Tadić et al., 2020). The lack of association in these studies is more likely due to the small sample size, which ranged from 27 to 150 participants, which reduces the power to detect the true effect.

5.5.1. limitations

This study also had some key limitations. Firstly, body composition differs between men and women, with men having more muscle mass and women having proportionally more fat mass (Schorr et al., 2018). Body composition is also affected by the ageing process, which is generally characterized by an increase in total body FM and a concomitant reduction in lean mass and bone density, which are independent of general and physiological changes in body weight and BMI (Ponti et al., 2020). These changes in body composition could induce differences in the causal relation between anthropometric-related measures and MS risk/severity in men and women of different ages. Due to the lack of GWAS results based on sex/age for MS, we were unable to predict such differences or determine which anthropometric-related measure could strongly and precisely predict the risk of MS developing or MS

disability worsening among obese men and women. Secondly, we could not conduct a bidirectional MR analysis between anthropometric measures and MS severity due to the absence of variants with strong evidence of association with MS severity. Therefore, we have not ruled out a possible bidirectional causal relationship between anthropometric measures and MS severity. Thirdly, the other important issue that might affect the anthropometric-MS severity association is collider bias. A collider is a variable that is influenced by two or more variables (Griffith et al., 2020). In MR of progression, collider bias is a critical issue and can occur when the studied sample is only restricted to cases, resulting in independent risk factors becoming spuriously associated in the cases (Paternoster et al., 2017) (**Figure 5.2**) . For example, MS risk becomes a collider variable because it is influenced by the obesity-related measures and other risk factors (e.g. HDL). Conditioning on MS risk (i.e., selecting case only individuals) will induce spurious associations between those obesity-related measures and the other risk factors.

Thus, in MR, when MS severity is considered as an outcome (i.e., including only MS cases), for any risk factor that causes MS risk, the genetic variants for that risk factor (e.g., BMI) may be spuriously associated with other risk factors for MS risk (Paternoster et al., 2017) and so, the association between those genetic variants and MS severity may be subject to confounding by these factors. As a consequences, the MR assumption that "the genetic instrument is independent of factors that confound the association of the exposure and the outcome" will be violated (Paternoster et al., 2017). Thereby caution is needed in interpreting the association between obesity-related measures and MS severity, as these relationships might be susceptible to collider bias. Fourthly, MR estimates reflect the lifelong effects of a risk factor in contrast to the short-term effects captured in observational studies.

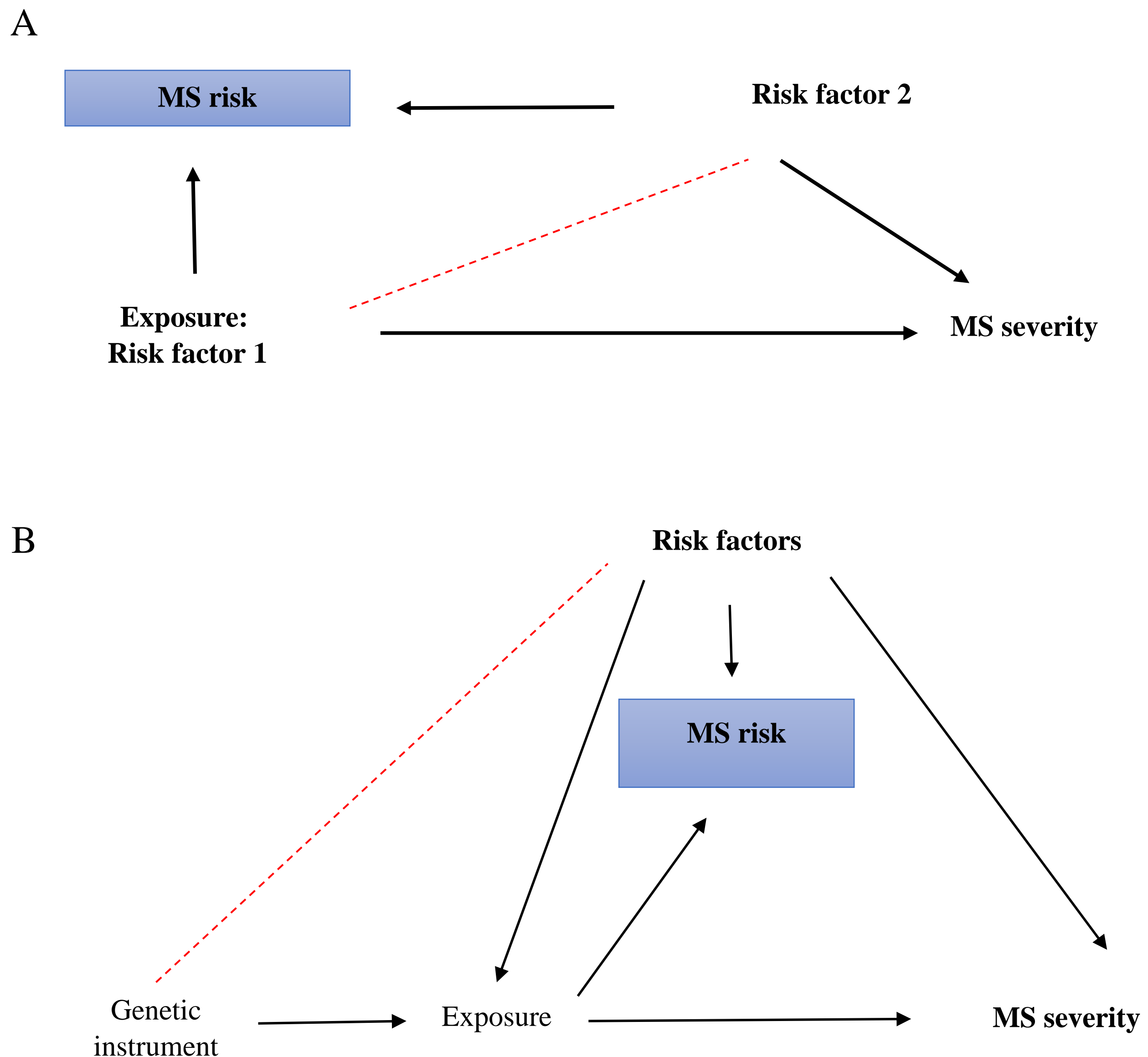


Figure 5.2: Direct acyclic graph Adapted collider bias in case-only studies. (A) MS risk becomes a collider variable because two arrows from risk factors collide at it. Consequently, spurious associations may exist between risk factors, both genetic and nongenetic, that are independently associated with MS risk. This opens up a noncausal pathway from exposure to MS severity when the risk factor is also associated with the MS severity. (B) In Mendelian randomization analysis, the genetic instruments cause the disease, and thus they become spuriously associated with independent risk factors, which are confounders of the exposure-MS severity. Conditioning on disease risk opened noncausal pathways between all variables that cause MS risk. This graph adapted from (Mitchell et al., 2018).

5.5.2. Conclusion

In conclusion, our findings provide evidence that obesity is an important contributor to MS development and MS severity, but height and non-fat mass are not. These findings expand our understanding of the role of anthropometric measures in MS aetiology. Importantly, these findings also identify a potentially modifiable factor that may reduce the accumulation of further disability and ameliorate MS severity.

Chapter 6 *Stroke as a risk factor for MS*

Statement of contribution

I designed this project, wrote and executed the analysis scripts used in this chapter myself. Nicholas W. Wood contributed to the interpretation of my results.

6.1. Introduction

Stroke and MS are common neurological diseases that are also common causes of disability worldwide (Hong et al., 2019). A stroke happens when there is a lack of oxygen and nutrients to brain tissue caused by an interruption to the blood flow, resulting in neuronal cell death and inflammation (Vijayan et al., 2017, Shi et al., 2019). There are two main types of stroke. Ischaemic stroke (IS) accounts for about 87% of all cases and occurs when a thrombus forms inside blood vessels or a blood clot migrates from the periphery, blocking the blood flow in a cerebral artery (Aloizou et al., 2021). IS can be further classified into large artery atherosclerotic stroke (LAS), cardioembolic stroke (CES) and stroke caused by small vessel stroke (SVS) (Vijayan et al., 2017, Malik et al., 2018). The second type of stroke is haemorrhagic stroke, which accounts for about 15% of cases and can be classified into two subtypes, intracerebral haemorrhage and subarachnoid haemorrhage (Vijayan et al., 2017).

Although the causes (ischaemia/thromboinflammation vs. autoimmunity) and clinical presentations of stroke and MS set them apart, they share common downstream mechanisms that can lead to both damage and recovery (Pinheiro et al., 2016, Aloizou et al., 2021). Demyelination and axonal injury are common features of MS, but are also found in stroke (Pinheiro et al., 2016), while vascular impairment and neurodegeneration are characteristics of stroke that are also observed in MS (Pinheiro et al., 2016). Interestingly, the most conspicuous common feature is the neuroinflammatory response, which is marked by glia cell activation and immune cell migration into the CNS (Pinheiro et al., 2016, Aloizou et al., 2021). Not surprisingly, the two diseases also share genes, gene expression levels and pathways (Tian et al., 2020a, Li et al., 2019).

Cumulative evidence suggests that patients with MS might have substantially higher risks of stroke than healthy people, but this relationship remains controversial. For example, a recent systematic meta-analysis that included more than 380,000 participants reported a greater risk of developing any type of stroke—and IS in particular—in people with MS than in non-MS populations during the first year after clinical onset (Hong et al., 2019). This finding is not surprising, given that inflammation is central to

the pathogenesis of MS, and it is possible to explain the increased incidence of stroke among patients with MS using the concept of inflammatory-driven atherosclerosis (Caprio et al., 2016). During inflammation, inflammatory cytokines increase inflammatory cell infiltration and oxidative stress, impair endothelial function, accelerate the formation and disruption of atherosclerotic plaques and thus increase the risks of cerebrovascular diseases, including stroke (Yuksel et al., 2019, Hong et al., 2019). However, a recent MR analysis identified no causal role for MS in stroke risk (Peng et al., 2022).

Nevertheless, because of the methodological limitations of conventional observational studies, the observed positive associations between MS and stroke might be due to confounding factors or reversed causality. For instance, there are several risk factors for IS, including blood pressure (Georgakis et al., 2020), HDL-C(Hindy et al., 2018), LDL-C (Hindy et al., 2018) and obesity (Mitchell et al., 2015), that are also established risk factors for MS. Such risk factors may therefore be potential confounds that produce spurious associations between stroke and MS, and none of the above studies has evaluated and eliminated the effects of these factors. These data indicate the need for more studies to determine the nature of the associations between MS and stroke using novel methods that can overcome the limitations of conventional observational studies such as MR.

6.2. Aims

Adjusting for confounds (blood pressure, HDL-C, LDL-C and obesity), this study aimed to examine the causal link between different kinds of stroke—including any stroke (AS), any IS and any IS subtype (LAS, CES and SVS)—and MS risk and severity by performing several MR analyses in order to evaluate the following relationships:

1. Causal effects of lifelong, genetically predicted MS on the risk of stroke.
2. Causal effects of lifelong, genetically predicted stroke on the risk of developing MS and on the severity of MS.

6.3. Method

Summary statistics for SNPs associated with MS risk, MS severity and the different types of stroke were derived from IMSGC and MEGASTROKE data (method section 2.2.3 and 2.2.2.4, respectively). Other GWAS datasets, including BMI, whole body FP and FM, blood pressure (systolic blood pressure (SBP) and diastolic blood pressure (DBP)), HDL-C, LDL-C and serum 25-hydroxyvitamin D levels, were obtained from MR-Base, which is an online database and analytical platform for MR analysis (www.mrbase.org) that allows users to download publicly available GWAS data (**Table 6.1**).

The SNPs significantly (p-value $< 5 \times 10^{-8}$ or p-value $< 5 \times 10^{-7}$) associated with exposure were clumped at an LD-clumping threshold of $r^2 = 0.01$. The Steiger filtering method was then applied, and mean F-statistics were calculated. For the MR analyses, IVW was used as the main estimator, while MR−Egger, weighted median and MVMR were used as sensitivity estimators. Heterogeneity across estimates was assessed using I^2 and Cochran's Q. The MR−Egger intercept was used to detect unbalanced horizontal pleiotropies. To correct for multiple tests, FDR was used to adjust the p-values computed for the IVW method to mitigate false-positive rates. Exposures with significant adjusted p-values of 0.05 or less were considered to be potential evidence of a causal effect.

Table 6.1: GWAS datasets included in the MVMR analysis

Accession ID	Trait		Population	Sample size	Author
ukb-a-359	Blood pressure	Diastolic blood pressure automated reading	European	317,756	Neale Lab
ukb-a-360		Systolic blood pressure automated reading		317,754	
ukb-a-248	Obesity	Body mass index (BMI)		336,107	
ukb-a-265		Whole body fat mass (FM)		330,762	
ukb-a-264		Body fat percentage (FP)		331,117	
ebi-a-GCST90000615		Serum 25-Hydroxyvitamin D levels	European	417,580	(Revez et al., 2020)
ieu-a-300	cholesterol	LDL-C	Mixed [a]	173,082	(Willer et al., 2013)
ieu-a-299		HDL-C		187,167	

[a] principal-components or mixed-model approaches were used to adjust for population structure in studies with mixed populations (i.e., European and non-European). Abbreviations: MVMR, multivariable mendelian randomization.

6.4. Results

6.4.1. No causal link between genetically predicted MS susceptibility and the risk of stroke

After evaluating LD and removing palindromic SNPs, 166−159 SNPs associated with MS risk were included in the MR analysis. The mean F-statistic for these SNPs was greater than 10, which suggests that weak instrument bias is likely negligible among these data. The main results are shown in Table 6.2, Figure 6.1 and Figure 6.2. The results for the IVW, MR–Egger and weighted median methods indicated no causal effect of MS on the risk of the different types of stroke (AS, IS, LAS, CES and SVS). The directions of the estimates of the associations of MS with different types of stroke using different MR methods were consistent, except for the association of MS with IS, in which MR–Egger showed the opposite direction to IVW and the weighted median, suggesting the presence of outliers. The heterogeneity test identified evidence of heterogeneity in the analysis of MS and IS, but not in the analysis of MS and other types of stroke. The MR–Egger intercept provided no evidence of pleiotropy.

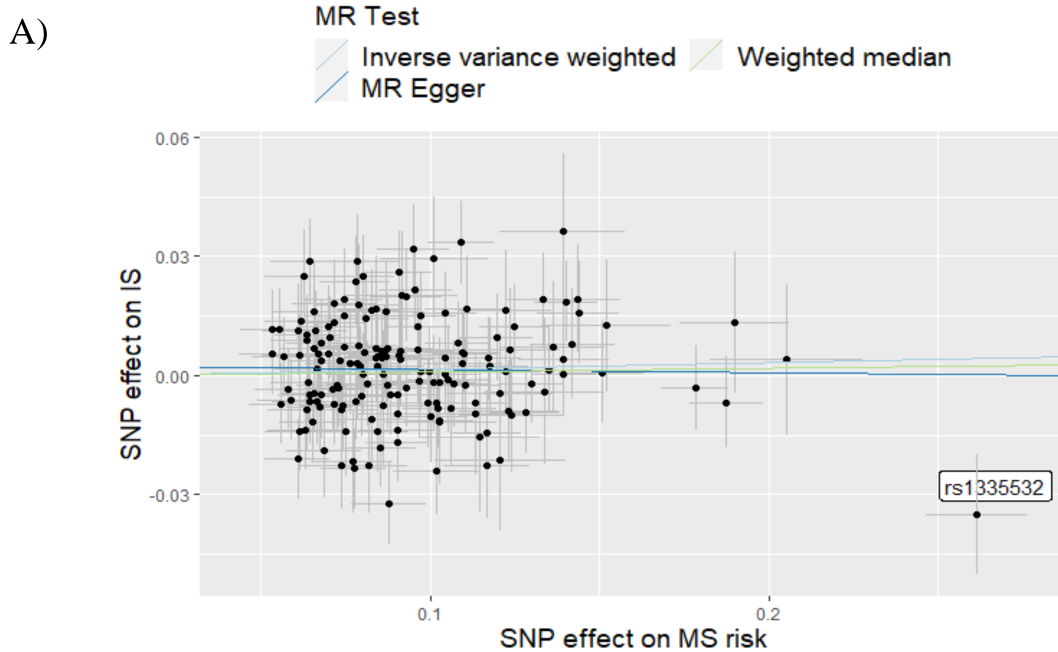
Inspection of the scatter plots highlighted one outlying SNP (rs1335532) responsible for the heterogeneity and the inverted estimate from MR–Egger in the analysis of MS and IS. After removing this outlier, the directions of the effect estimates were consistent among the MR methods, and the heterogeneity was reduced but still significant. The MR causal effect estimates remained non-significant, further confirming the absence of a causal effect of MS on IS.

Table 6.2: The estimated effects of MS on the risk of strokes based on MR

Traits		No. of SNP	F test	MR estimates OR (95%)							Pleiotropic assessment				
				IVW	P value	FDR	MR Egger	P value	Weighted median	P value	Q	Q P value	I^2 (%)	MR-Egger intercept	MR-Egger intercept P value
AS	All SNPs	164	71	1.02(1,1.03)	9.49E-02	2.36E-01	1.03(0.96,1.1)	4.65E-01	1.01(0.99,1.04)	3.34E-01	190	6.70E-02	14.6	-0.0009	7.81E-01
IS	All SNPs	166	72	1.02(1,1.04)	1.18E-01	2.36E-01	0.99(0.92,1.06)	8.21E-01	1.01(0.98,1.04)	5.39E-01	208	1.11E-02	21.1	0.0024	4.74E-01
	No outliers	165		1.02(1,1.04)	5.12E-02		1.02(0.95,1.1)	5.47E-01	1.01(0.98,1.04)	4.78E-01	202	2.08E-02	19.2	-0.0002	9.48E-01
LAS	All SNPs	159	70	0.97(0.93,1.02)	2.80E-01	2.80E-01	0.89(0.75,1.06)	1.94E-01	0.97(0.9,1.03)	3.23E-01	160	4.11E-01	2.1	0.0087	2.94E-01
SVS	All SNPs	164	70	1.03(0.99,1.07)	1.89E-01	2.36E-01	1.03(0.87,1.21)	7.57E-01	1.02(0.96,1.09)	4.71E-01	163	4.58E-01	0.8	0.0003	9.64E-01
CES	All SNPs	164	70	1.03(0.99,1.07)	1.69E-01	2.36E-01	1.12(0.97,1.28)	1.21E-01	1.06(1,1.12)	4.10E-02	175	2.35E-01	7.3	-0.008	2.20E-01

Abbreviations: No. of SNPs, number of single-nucleotide polymorphisms; F test, mean F-statistic; IVW, inverse-variance weighted; OR, odds ratio; CI, confidence interval; I^2 (%), the level of heterogeneity expressed as a percentage; Q, Cochran's Q test; AS, any stroke; IS, ischaemic stroke; LAS, large artery atherosclerotic stroke; SVS; small vessel stroke; CES, cardioembolic stroke.

A)



B)

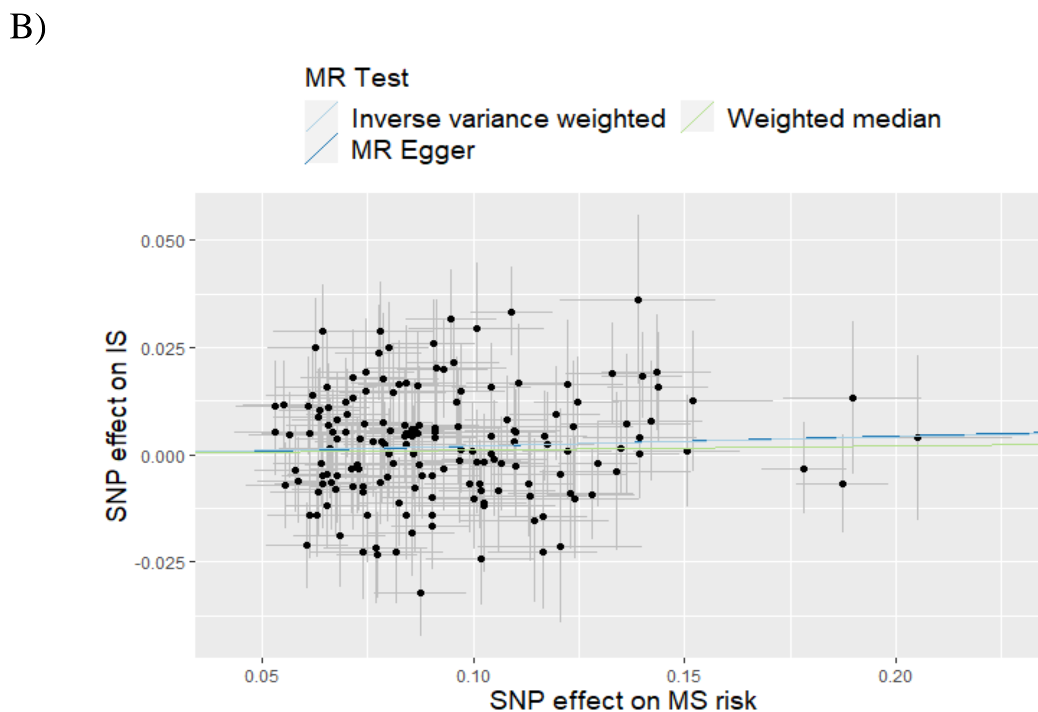


Figure 6.1: **(A)** a scatter plot relating the effect sizes (beta-coefficients, black points) of the SNP–MS risk associations and the SNP–IS risk associations with 95% confidence intervals (grey lines). The slopes of the lines represent the causal estimates using each of the three different methods. Outlier SNPs are labelled**. (B)** the scatter plot after excluding the outlier SNPs. Abbreviations: IS, ischaemic stroke; MS, multiple sclerosis.

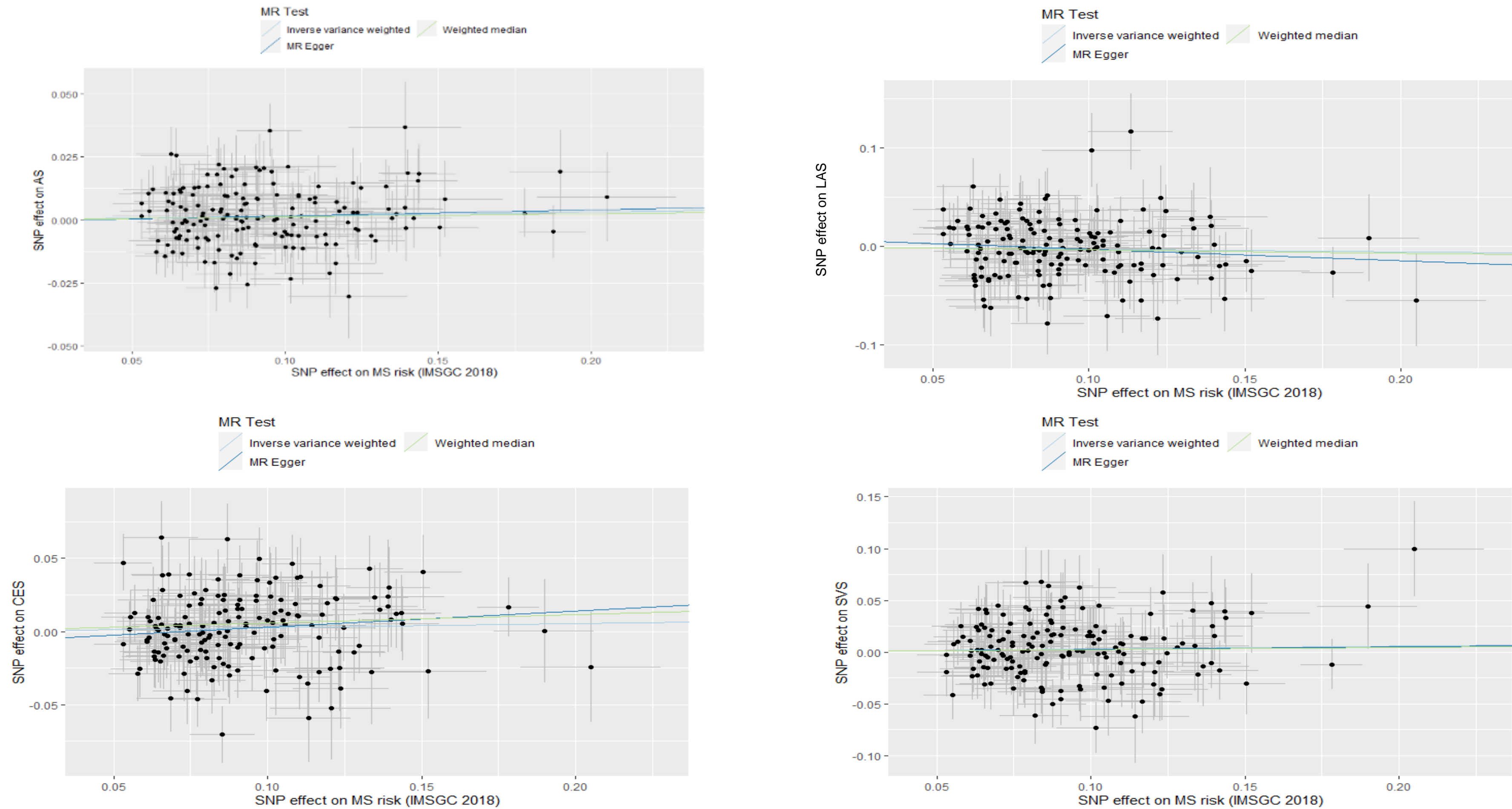


Figure 6.2: A scatter plot relating the effect sizes (beta-coefficients, black points) of the SNP–IS associations and the SNP–MS risk associations with 95% confidence intervals (grey lines). The slopes of the lines represent the causal estimates using each of the three different methods. Outlier SNPs are labelled**.** Abbreviations: AS, any stroke; LAS, large artery atherosclerotic stroke; SVS, small vessel stroke; CES, cardioembolic stroke; MS, multiple sclerosis.

6.4.2. Causal link between genetically predicted IS and MS risk

To examine the causal effect of different types of stroke on the risk of MS developing, 61 SNPs strongly and independently associated with stroke (23 SNPs associated with AS, 15 with IS, nine with CES, eight with LAS and six with SVS) were included in the MR analyses. The mean F-statistic for these SNPs was greater than 10, suggesting that weak instrument bias was likely negligible among these data. The main results are shown in Table 6.3, Figure 6.3 and Figure 6.4. The MR results from IVW demonstrated that genetically predicted IS was causally associated with an increased risk of developing MS. The results from the weighted median were comparable in both direction and effect size and were statistically significant, further supporting the IVW results, although MR–Egger produced a non-significant estimate due to a wider CI. There were no causal effects of AS, LAS, CES or SVS on MS risk.

The direction of the estimates for the associations of IS, LAS and SVS with MS risk was consistent between IVW and the weighted median, but MR–Egger had an opposing direction of association, although the direction was consistent for the association of AS and CES with MS risk. The heterogeneity test was significant for the analyses of stroke and MS risk, as shown by Cochran's Q test, except in the analysis of LAS and MS risk. However, the MR–Egger intercept estimate provided no evidence for pleiotropy in any of the analyses.

Inspection of the scatter plots highlighted three outlier SNPs (rs11242678, rs4942561 and rs60102266) for IS, three (rs35818742, rs7766042 and rs76576182) for SVS, one (rs114279112) for CES and three (rs10774624, rs11242678 and rs7219031) for AS. These outliers were found to be responsible for the excessive heterogeneity and for the inverted MR–Egger estimate. To determine the impact of the outliers on the causal estimates, they were removed and the MR analysis repeated. The heterogeneity then became non-significant, and the direction of the estimates from all MR methods became consistent. The estimated effects remained significant for the association of IS with MS risk and non-significant for the associations of AS, LAS and CES with MS risk. Interestingly, after removing the

outlier SNPs, the results provided evidence supporting the causal effect of SVS on MS risk such that a one-unit increase in the genetically predicted log-transformed odds of SVS could increase the risk of MS.

To further validate the causal effects of IS on MS risk, MVMR analysis was performed to avoid pleiotropy. The analysis indicated that, after adjusting for cholesterol, obesity, blood pressure and serum 25-hydroxyvitamin D levels, the causal effects remained statistically significant (OR = 1.53, 95% CI: 1.36– 1.72, p = 7.82E-13), suggesting the causal effect of IS on MS risk cannot be explained by these confounding factors and that IS acts as an independent risk factor for MS.

Table 6.3: The estimated effects of strokes on the risk of MS based on MR

Traits		No. of SNP	F test	MR estimates OR (95%)							Pleiotropic assessment				
				IVW	P value	FDR	MR Egger	P value	Weighted median	P value	Q	Q P value	I^2 (%)	MR-Egger intercept	MR-Egger intercept p value
AS	All SNPs	23	32	1.1(0.89,1.37)	3.57E-01	6.25E-01	1.76(0.43,7.25)	4.40E-01	1.12(0.91,1.38)	2.75E-01	58	2.73E-05	63.7	-0.0288	5.19E-01
	No outliers	20		1.15(0.97,1.35)	1.01E-01		1.64(0.59,4.51)	3.54E-01	1.13(0.92,1.38)	2.44E-01	25	1.33E-01	27.2	-0.0218	4.94E-01
IS	All SNPs	15	35	1.37 (1.08,1.73)	**9.12E-03**	**2.13E-02**	0.7 (0.21,2.32)	5.72E-01	1.25 (0.99,1.56)	**5.65E-02**	36	6.68E-04	63.6	0.0467	2.87E-01
	No outliers	12		1.6 (1.32,1.93)	**1.18E-06**	**8.26E-06**	1.11(0.4,3.06)	8.46E-01	1.42 (1.11,1.82)	**4.97E-03**	15	1.21E-01	34.6	0.0251	4.90E-01
LAS	All SNPs	8	32	1.01(0.92,1.1)	9.08E-01	9.42E-01	0.87(0.67,1.13)	3.36E-01	1.02(0.9,1.15)	7.61E-01	2	9.14E-01	0	0.0338	2.87E-01
	No outliers	5		1.05(0.94,1.18)	3.91E-01		1.06(0.6,1.88)	8.52E-01	1.08(0.95,1.23)	2.58E-01	1	7.32E-01	0	-0.002	9.75E-01
SVS	All SNPs	6	27	1.1(0.82,1.47)	5.18E-01	7.25E-01	0.54(0.14,2.1)	4.21E-01	1.07(0.87,1.32)	5.37E-01	23	1.07E-04	82.9	0.1244	3.51E-01
	No outliers	3		1.41(1.14,1.76)	**1.89E-03**	**6.62E-03**	1.99(0.36,10.86)	5.73E-01	1.41(1.12,1.79)	3.93E-03	3	6.38E-02	70.9	-0.0541	7.57E-01
CES	All SNPs	9	76	1(0.89,1.12)	9.42E-01	9.42E-01	1.1(0.83,1.46)	5.28E-01	1.07(0.96,1.19)	2.04E-01	15	4.03E-02	52.3	-0.0165	4.66E-01
	No outliers	8		1.01(0.9,1.14)	8.21E-01		1.21(0.92,1.59)	2.28E-01	1.08(0.96,1.21)	1.86E-01	10	1.16E-01	41.2	-0.0279	2.22E-01

Abbreviations: No. of SNPs, number of single-nucleotide polymorphisms; F test, mean F-statistic; IVW, inverse-variance weighted; OR, odds ratio; CI, confidence interval; I^2 (%), level of heterogeneity expressed as a percentage; Q, Cochran's Q test; AS, any stroke; IS, ischaemic stroke; LAS, large artery atherosclerotic stroke; SVS; small vessel stroke; CES, cardioembolic stroke.

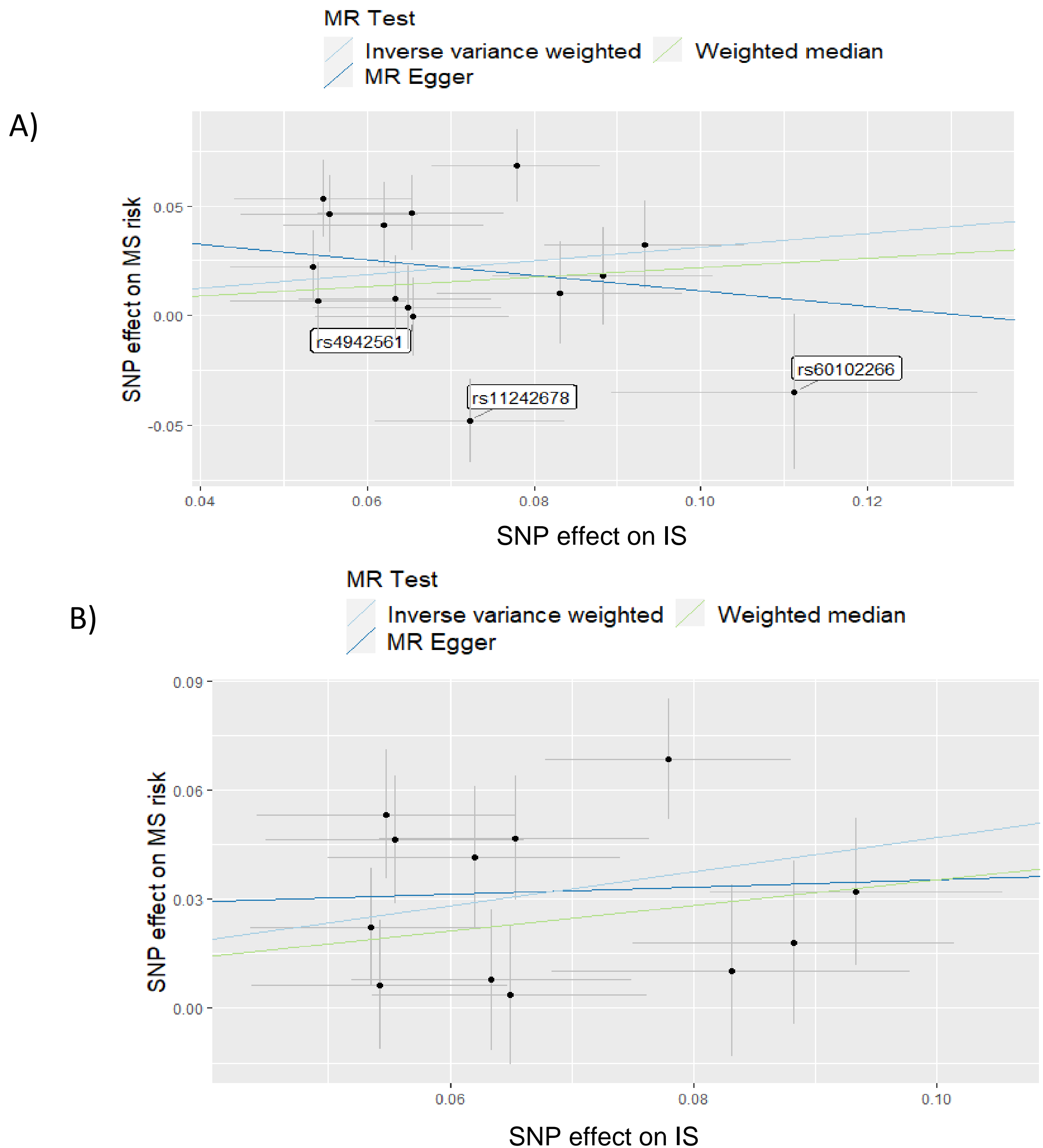


Figure 6.3: **(A)** a scatter plot displaying the effect sizes (beta-coefficients, black points) of the SNP–IS association and the SNP–MS risk associations with 95% confidence intervals (grey lines). The slopes of the lines represent the causal estimates using each of the three different methods. Outlier SNPs are labelled**. (B)** the scatter plot after excluding the outlier SNPs. Abbreviations: IS, ischaemic stroke; MS, multiple sclerosis.

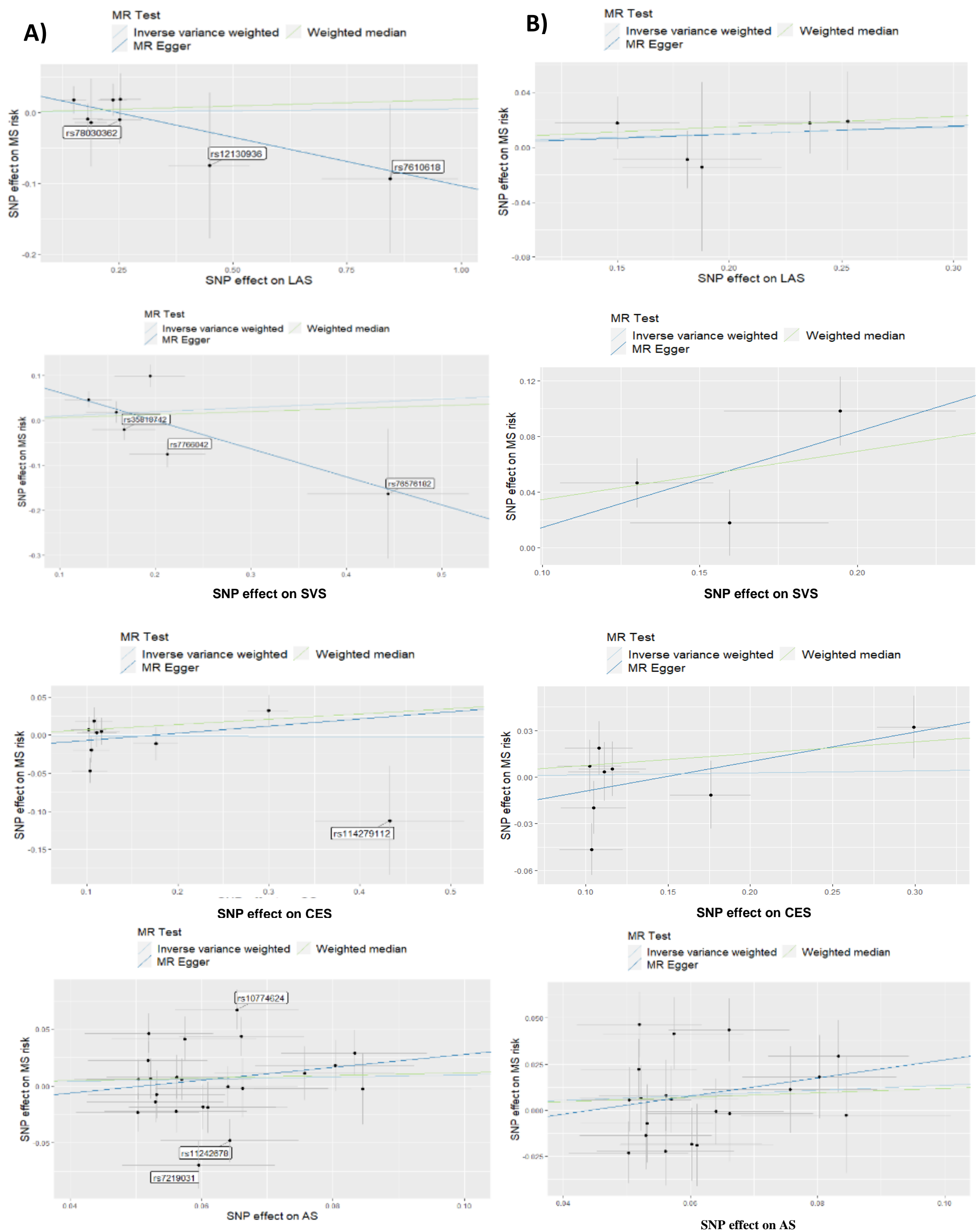


Figure 6.4: (A) scatter plots displaying the effect sizes (beta-coefficients, black points) of the SNP–AS and SNP–IS-subtype associations and the SNP–MS risk association with 95% confidence intervals (grey lines). The slopes of the lines represent the causal estimates using each of the three different methods. Outlier SNPs are labelled**. (B)** the scatter plots after excluding the outlier SNPs.

Abbreviations: LAS, large artery atherosclerotic stroke; SVS, small vessel stroke; CS, cardioembolic stroke; AS, any stroke; MS, multiple sclerosis.

6.4.3. No causal link between genetically predicted stroke and MS severity

To examine the causal effect of stroke on MS severity, one SNP for SVS, five for AS and two each for IS, LAS and CES were included in the MR analyses. The mean F-statistic for these SNPs was greater than 10, suggesting that weak instrument bias was likely negligible among these data. The main results are shown in Table 6.4. The MR analysis indicated no causal effect on MS severity for any type of stroke. No evidence for heterogeneity was identified in the analyses. The MR–Egger intercept test was performed only for the AS analysis because the number of SNPs was greater than three. The results indicated no pleiotropy in the association between AS and MS severity.

Table 6.4: The estimated effects of stroke on MS severity based on MR

Trait	Method	NO. of SNPs	F test	MR estimates OR (95% CI)			Pleiotropic assessment				
				Beta (95% CI)	P value	FDR	Q	Q P value	I^2 (%)	MR-Egger Intercept	MR-Egger Intercept P value
IS	IVW	2	31	0.65 (-0.35, 1.65)	2.01E-01	6.39E-01	0.3	5.60E-01	0		
LAS	IVW	2	36	-0.25 (-1.01,0.501)	5.11E-01	6.39E-01	3.3	6.00E-02	0		
SVS	Wald ratio	1	24	0.308 (-0.49, 1.11)	4.50E-01	6.39E-01					
CES	IVW	2	108	-0.032 (-1.96, 1.96)	8.27E-01	8.27E-01	1.02	3.00E-01	0		
AS	IVW	5	36	0.29 (-0.35,0.93)	3.74E-01	6.39E-01					
	MR Egger	5		0.24 (-3.17,3.64)	9.01E-01		0.236	9.70E-01	0	0.0042	9.76E-01
	Weighted median	5		0.34 (-0.41,1.1)	3.74E-01						

Abbreviations: No. of SNPs, number of single-nucleotide polymorphisms; F test, mean F-statistic; IVW, inverse-variance weighted; OR, odds ratio; CI, confidence interval; I^2 (%), the level of heterogeneity expressed as a percentage; Q, Cochran's Q test; AS, any stroke; IS, ischaemic stroke; LAS, large artery atherosclerotic stroke; SVS; small vessel stroke; CES, cardioembolic stroke.

6.5. Discussion

In this study, MR analyses were performed to investigate a potential causal link between stroke and its subtypes and MS (risk and severity). The results showed that 1) genetically predicted IS was associated with increased risk of MS and that this relationship was independent of IS–MS confounds; 2) MR analysis of IS subtypes provides genetic support for a causal effect of SVS on MS risk; 3) evidence for the causal role of stroke in MS severity remains uncertain; and 4) genetically predicted MS has no causal role in the risk of stroke.

Cumulative evidence shows that, before the clinical onset of MS, patients with MS had a decreased risk of IS, but a year after clinical onset, they had an increased risk (Thormann et al., 2016, Zöller et al., 2012, Tseng et al., 2015), indicating that the longer an individual is exposed to MS, the higher the excess risk for IS (Thormann et al., 2016), perhaps due to longer periods of inflammation. Since the SNPs−MS were associated with loci participating in biological mechanisms leading to the risk of MS, not the disease progression this may explain the absence of the causal role of MS on stroke in this current study.

In line with the current findings, a recent MR analysis that used 44 SNPs extracted from the discovery phase of 2013 MS risk GWAS (IMSGC) identified no causal association between MS and the risk of AS (Peng et al., 2022). However, in that study, no MR analyses were performed to examine the causal role of MS on other types of stroke or the causal role of stroke in MS risk (Peng et al., 2022). Several other MR analyses were therefore conducted in the current study to explore the causal role of different types of stroke in MS.

Interestingly, the MR analyses indicated a causal effect of IS on MS risk, and this causal relationship appears independent of IS–MS confounds, such as cholesterol, blood pressure, obesity and serum 25-hydroxyvitamin D levels. Further subgroup analysis indicated a causal effect of SVS on MS risk after excluding outlier SNPs. No causal effects of other strokes on MS risk were identified. The lack of

evidence for a causal relationship between AS, LAS, CES and MS risk is unexpected due to a power issue because SNP–AS, SNP–LAS and SNP–CES explained 0.9%, 1.8% and 1.4%, respectively, of the variance in these phonotypes—relatively more than the variance explained by the SNPs in IS (0.9%) and SVS (0.8%).

Although MS tends to affect young adults between the ages of 20 and 40 years and strokes tend to affect people 65 or older (Yousufuddin and Young, 2019, Naseri et al., 2021), both diseases can arise at any age. The incidence of late-onset MS has been examined in several studies that have reported patients experiencing their first symptoms at the age of 60 or older (Martinelli et al., 2004), including a case report of MS onset in an 82-year-old (Martinelli et al., 2004). Likewise, data on incidence and prevalence indicate that IS is no longer a disease affecting just elderly people, with 10% to 20% of IS events occurring in people aged 18 to 50 years (Boot et al., 2020). In the current analysis, the ages of cases of IS and SVS GWASs range from 15 to 65 years old, and the age for MS cases range from 18 to 69 years old, which provides an overall idea of how IS and SVS may happen first.

It is unclear which biological mechanisms could explain the causal effect of IS on MS risk, but recent evidence suggests the possible role of autoimmunity after strokes. IS causes damage to neurons, glia and the vasculature, resulting in BBB damage, haemorrhage, oedema and necrotic cell death (Javidi and Magnus, 2019). Dying and dead cells prompt the release of danger signals that activate the immune system (Javidi and Magnus, 2019). The sterile inflammatory reaction then involves the innate immune system, with the activation of the resident immune cells of the CNS, such as microglia, and a rapid infiltration of peripheral immune cells, including neutrophils, dendritic cells, macrophages and lymphocytes, into the brain through the compromised BBB (Javidi and Magnus, 2019). After the lymphocytes infiltrate the ischaemic brain and brain antigens are released into the CNS and peripheral circulation, IS may induce a secondary autoimmune response due to direct contact between brain antigens and lymphocytes in the CNS and peripheral circulation (Javidi and Magnus, 2019).

Indeed, 24 hours following a stroke, a higher peak concentration of autoantigens (i.e., self-antigens) such as MBP, neuron-specific enolase (NSE) and S100beta, has been observed in the serum of patients with stroke (Jauch et al., 2006). These antigens can be internalised by macrophages and other antigen-presenting cells and present on MHC class (Wang et al., 1992). The antigens bound to MHC class II molecules can be recognised by T-cells, which may then be activated, proliferate and/or secrete cytokines, such as interferon-gamma (Wang et al., 1992). High numbers of T-cells secreting interferon-gamma in response to MBP and myelin proteolipid protein (PLP) have been observed in patients with IS (Wang et al., 1992). Indeed, a damaged BBB, activated microglia, T-cells infiltrating into the brain and the damaging effect of adaptive immunity are well established in MS patients and MS animal models (EAE and autoimmune demyelination) (Javidi and Magnus, 2019, Iadecola and Anrather, 2011, Schaeffer et al., 2015b), and at a biological level, this mechanism may therefore offer the most intriguing hypothesis to explain the causal role of IS in MS risk.

Although there is an accumulation of evidence that describes IS risk in people with MS (Hong et al., 2019), MS incidence in IS cases remains unknown—exacerbation of MS after stroke has been described, but it seems rare. So far, the studies undertaken have presented preliminary results in an abstract form, but final results have not been published. For example, a rare case was reported of a female patient with SPMS who developed high disease activity (relapse) one month after IS of the right middle cerebral artery (Poellmann et al., 2015). The authors suggested that breaking the BBB after stroke with an influx of activated immune cells into the CNS could explain MS exacerbation after stroke (Poellmann et al., 2015), which may support the hypothesis proposed above.

Recent evidence indicates that IS and MS are genetically linked, with several shared genes, gene expression levels and pathways (Li et al., 2019, Tian et al., 2020b). Several loci associated with IS in the current MR analysis, have also been reported in MS. These loci participate in immunity and inflammation processes, suggesting distinct pathogenesis mechanisms for the development from IS to MS.

For example, *SH2B3* is an adaptor protein that functions as a negative regulator of cytokine signalling pathways (Flister et al., 2015). The rs3184504 [T] allele of *SH2B3* causes a missense mutation associated with leucocytosis, enhanced innate immunity and increased susceptibility to cardiovascular and autoimmune disorders (Flister et al., 2015), including MS (Alcina et al., 2010).

In addition, the rs1975161 variant falls in the downstream region of ILF3, which forms a heterodimer with ILF2 that is required for the T-cell expression of IL-2, a proinflammatory cytokine that regulates various aspects of inflammation (Malik et al., 2018, Beecham et al., 2013b). IL-2 and other cytokines, including IFN-γ and TNF-α, are widely believed to initiate MS and contribute to oligodendrocyte damage progression (Nasl-khameneh et al., 2018). Furthermore, rs42039 is located in the 3'-untranslated regions of CDK6, which is recognised as a novel kinase phosphorylating nuclear factor kappa B (NF-kB), an important contributor to inflammation in MS (Zhou et al., 2020).

As with MS risk, the causal relationship between strokes and MS severity was also explored in this study, but the MR analyses found no evidence for such a relationship. This lack of evidence may reflect insufficient statistical power due to the small number of SNPs included in the MR analyses of MS severity.

6.5.1. Limitation

There are several advantages of the current study. First, this appears to be the first comprehensive MR analysis that has examined the causal link between different types of stroke and MS risk and severity. Second, the MR analyses were conducted with all SNPs and without outliers to ensure the validity of the results and to mitigate outlier effects. Third, MVMR was employed to guard against confounders caused by pleiotropy and to produce robust MR results. Finally, the summary statistics for the SNPs were extracted from a large-scale GWAS of MS and stroke.

Nevertheless, the analysis also has limitations. The current evidence in Chapter 6 might not be enough to prove that ischemic stroke (IS) is a risk factor for MS. This is because I was unable to find any reports on the MS incidence in IS cases, and, so far, the studies undertaken have presented only preliminary results in an abstract form, with final results having not yet been published. Although this finding may be considered novel—especially nowadays, as there are several reports suggesting that MS and stroke, specifically IS, share common pathways and genes—care must be taken when interpreting significant results in this analysis. This is because, in some MR analyses, MR–Egger had an opposing direction of association, indicating the presence of outliers. After outlier removal, the directions of the MR estimates were consistent among all the methods (IVW, MR–Egger, and weighted median) and the results remained the same except for SVS. In the SVS analysis, the MR finding with outliers was statistically non-significant, but, after removing the outliers (3 SNPs out of 6), it became statistically significant. Although outlier removal methods can effectively reduce bias in MR estimates, caution must be taken to avoid overinterpretation, because this strategy will reduce the standard error of the causal effect estimate and might lead to over-fitting (Hemani et al., 2018a). In such a case, replicating the finding would be an effective strategy to ensure the reliability of the findings. Due to the lack of appropriate independent stroke data, I could not conduct a replication analysis. Furthermore, although METASTROKE is the most large-scale GWAS meta-analysis study, this study reported a small number of SNPs that reached genome-wide significance levels: 18 loci (12 novel) for AS, 20 (12 novel) for IS, 6 (3 novel) for LAS, 4 (2 novel) for CES, and 2 novels for SVS. In this regard, it has been noted that other genetic variants associated with stroke in other studies might not be captured to some extent. Therefore, this MR should be explored in-depth following the availability of larger genetic data in the future Another limitation is that the MR analyses were performed using European ancestry, weakening the generalisability to other ancestries.

Although several confounding factors in the relationship between ischemic stroke and MS risk were controlled for in the MVMR model, the effect of relevant but unobserved potential confounders could not be excluded, which is an established limitation of the MR approach (Marini et al., 2020, Lee et al., 2020b).

Although MS tends to affect young adults between the ages of 20 and 40 years and strokes tend to affect people > 65, the findings indicate that IS is a risk factor for MS. This finding suggests the possibility that an individual who had IS at an early age may have an increased chance of having MS. To test this hypothesis, I sought to apply MR using genetic variants associated with early-onset stroke to assess of the effect of age on MS risk development. However, I failed to conduct such an analysis due to a lack of appropriate data.

6.5.2. Conclusion

In summary, MR analysis provided genetic evidence that IS and SVS, but not AS, LAS or CES, have a causal role in the increased risk of developing MS. No genetic evidence for causality between stroke and MS severity was found, and MR analysis also found no causal effect of MS on strokes. These findings shed light on IS as a risk factor for MS and, most importantly, suggest closer attention be paid to MS prevention for IS patients.

Chapter 7 Discussion

7.1. Highlighting the key limitations of genetic and non-genetic studies on informed causal inference and drug repurposing

The term "determinants of health" was first introduced in the 1970s, and it refers to several factors, including genetic, biological and lifestyle, that have a significant influence, whether positive or negative, on health (Marans, 2003). Epidemiologic studies played a unique role in identifying such health determinants and provided insight into the guiding diagnosis, therapy or disease control. For example, through GWAS, researchers have identified tens of thousands of common genetic variants associated with biomarkers, lifestyle factors and disease outcomes, including MS risk. However, understanding how these variants mechanistically influence disease phenotypes and/or translating GWAS findings into drug targets have proved challenging for several reasons (Jacobs et al., 2020). The majority of GWAS-identified SNPs are assumed not to be causal but to highlight a region of LD containing one or more functional SNPs (MacArthur et al., 2017). Linking the effect of SNPs to gene function is not straightforward without additional data, particularly as many of these disease-associated SNPs fall within non-coding regions of the genome with no direct impact on protein structure or function (MacArthur et al., 2017, Porcu et al., 2019). Observational studies have also contributed to identifying non-genetic health determinants such as biomarkers (e.g., lipids) or lifestyle factors (e.g., smoking). However, it has been unclear whether the factors identified by these studies were causal and contributed to disease pathogenesis or were simply due to confounding and/or reverse causation. Randomised clinical intervention trials are able to infer causality, but these studies are expensive and often have a limited duration of intervention trials, and information on potential long-term side effects may not be obtained (Benn and Nordestgaard, 2018).

7.2. Mendelian randomisation

The challenges facing epidemiological studies in distinguishing between causation and association have drawn much interest to the MR approach. MR can be viewed as a platform to integrate novel genetic information generated in GWAS or molecular studies (QTL) to inform about causal associations between an exposure and a disease outcome; the demonstrated causal association is less likely to be affected by confounding or reverse causation.

Given that MS is a multifactorial disease influenced by both genetic and environmental factors, demonstrating causation would aid the understanding of the role of these factors in MS pathogenesis, identify opportunities for prevention, and inform regarding potential drug targets. Therefore, this thesis set out to use genetics and the MR approach to address two critical questions:

- Whether the associations between a genetically predicted risk factor

and MS risk and severity, respectively, are causal.

- Whether the associations between the genetically predicted expression of genes that encode protein targets for approved drugs or drugs in clinical development and MS risk and severity, respectively, are causal.

As each results chapter included a discussion, I here briefly describe how I tackled the MR limitations. Following this, I summarise the main findings of this thesis.

7.3. Pipeline for Mendelian randomisation analysis

Several potential limitations apply to the MR design, the most notable being weak instruments, pleiotropy of instruments, and LD between genetic variants used as instruments (Benn and Nordestgaard, 2018). I sought to avoid such limitations by restricting the selection of instruments to SNPs associated with the exposure of interest at genome-wide significance (p-value 5×10^{-8}) and then estimated the mean F-statistics to avoid weak instruments.

The selected instruments were clumped to avoid LD between instruments. Steiger filtering was used to exclude SNPs that explained more variation in the outcome than variation in the exposure to avoid pleiotropic SNPs. I used different methods that were able to identify and correct for pleiotropy, including the MR-Egger intercept test, Cochran's Q test, I^2 test, MR-Egger, weighted median and MVMR. In addition, I attempted to identify invalid and/or outlier SNPs via the radia-MR test or scatter plot. These SNPs were removed, and the MR analyses were repeated to ensure the validity of the results. Furthermore, in the case of the correlated instrument ($r^2 > 0.2/0.4$), MR methods that take into account the correlation were used.

7.4. Mendelian randomisation for drug target repurposing

In Chapters Three and Four, my thesis discussed identifying potential drug repurposing opportunities for MS. I investigated the role of statins as therapeutic targets for MS risk and severity, respectively. I used a set of genes to mimic the effect of statins' cholesterol-dependent and cholesterol-independent pathways. This novel approach predicted a causal association between SNPs in the *RAC2* gene region and MS risk, suggesting that statins may reduce MS risk through an independent cholesterol pathway involving *RAC2*, a member of the family of Rho GTPases that represents a pleiotropic effect of statin therapy. However, no evidence of a causal relationship between statins and MS severity was observed. In Chapter Four, I sought to increase the opportunities for identifying therapeutic targets relevant to MS risk and severity by performing MR of the druggable genome. This comprehensive work revealed strong genetic evidence for the causal association between CCR4, SIK3 and SLAMF7 and MS risk, suggesting that the drugs target these genes may represent a repurposing opportunity for MS prevention. There is no evidence for the causal role of any of these druggable genes in MS severity.

7.5. Mendelian randomisation for causal analysis of risk factors

In Chapters Three, Five and Six, I described the use of SNPs to investigate the role of several risk factors in MS risk and severity, respectively. Using MR, I found that SNPs that act as proxies for 21 anthropometric-related measures yielded evidence to indicate that fat (obesity) is a causal risk factor for MS, but height and non-fat tissue are not. The remarkable aspect of Chapter Five was that I provided genetic evidence that fat-related measures were also a risk factor for worsening MS severity but height and non-fat tissue were not. From a clinical perspective, these findings highlight fat as a critical aetiology for both MS risk and severity and indicate the need for strategies to control body composition, which may have important implications for managing MS incidence and disability accumulation. In Chapter Six, I conducted several MR analyses to explore the causal link between different kinds of strokes and MS risk and severity, respectively. The results revealed a significant link between ischemic stroke and its subtype, small vessel stroke, and MS risk but not MS severity. The findings also indicate the absence of association between other types of stroke, including AS, LAS and CES, and MS outcomes. Interestingly, these results shed light on the possible critical role of IS in increasing the risk of developing MS. Further, I evaluated the causal associations between lipid fractions, including HDL-C, LDL-C and TG, and MS risk and severity. These lipid fractions were assessed in Chapter Three, with LDL-C assessed under the effect of the cholesterol-dependent pathway for statins. The results provide strong evidence that HDL-C is associated with an increased risk of MS, whereas LDL-C and TG may not be related to MS risk. No evidence was identified for the causal role of lipid fractions in MS severity. Given that only fat-related measures were linked to MS severity among the risk factors examined here, this finding indicates that a risk factor predisposing to MS's development is not necessarily related to MS severity. However, such findings may also imply a lack of statistical power as observed in Chapter six regarding strokes and MS severity.

In all the previous chapters, I have also examined the causal role of increased genetic liability to MS on anthropometric-related measures, strokes and lipid fractions. The results did not support a causal association between MS genetic liability and these exposures.

7.6. Future aims for MS severity and clinical courses

In the past few years, the study of genetics has been revolutionised by the rapid progress in GWASs that has led to uncovering a large proportion of the genetic variants that are concerned with the incidence of disease. In MS, IMSGC has identified to date more than 200 genetic loci, and these discoveries provide important biological insights into the genes and pathways associated with the development of MS. In addition, these GWASs helped to conduct different well-powered analytic studies, such as MR.

However, there is a general lack of genetic variants concerning disease progression and severity. Paternoster et al. reported that only a small proportion of GWAS studies [approximately 8% of associations curated in the GWAS Catalog (p-value $< 1 \times 10^{-5}$)] have attempted to identify variants associated with disease progression or severity, and most of those have a small sample size (90% have sample size < 5000) (Walker et al., 2017). Similarly, MS severity and different clinical MS courses (RR, SP and PP) have received less attention in genetic studies. As I mentioned in Chapter One, there are only a few GWASs for MS severity to date, and no single result has achieved genome-wide significance in these studies. This highlights the need for larger cohorts to identify genetic variants relating to disease severity and progression to be able to conduct well-powered studies to predict new drugs or unintended drug effects related to the treatment of MS. Such studies would transform clinical care, providing insights into the strategies needed to develop better diagnosis, prognosis and prevention and to guide effective treatments.

7.7. Future directions of Mendelian randomisation

Since MR was initially developed, it has been applied to a wide range of research areas, including drug target validation and prioritisation and the interpretation of multi-dimensional omics data (Zheng et al., 2019). There are still areas of ongoing methodological research in MR, but work is required to focus on areas related to MR limitations and methods.

MR studies have mainly focused on European-ancestry populations due to the availability of large-scale GWASs in European populations; however, the MR findings cannot be generalised to other ancestries. Therefore, increasing the availability of new exposure variables, the detail of genetic measurements in different ancestries, and the amount of publicly available data resources would make it feasible to conduct comprehensive MR studies with different ancestries.

GWASs of omics data, such as gene expression, protein expression, metabolites, and DNA methylation, provide the opportunity to infer the cause–effect relationships between thousands of molecular phenotypes and outcomes of interest in a MR framework. In such data, there are hundreds of SNPs, a majority of which are highly correlated in a single gene region and any of which could be used to assess the causal relationship; however, using too many of these SNPs in the analysis can lead to spurious estimates and inflated Type 1 error rates (Burgess et al., 2017c). On the other hand, using only a few SNPs can lead to ignoring the majority of the data, with estimates highly sensitive to the particular choice of SNPs (Burgess et al., 2017c). Thus far, it is not clear how to choose which SNPs to include in the analysis or what represents the best LD clumping threshold to obtain the most efficient estimate possible without the analysis suffering from numerical instabilities when there are large numbers of highly correlated candidate variants (Burgess et al., 2017c). In this regard, to the best of my knowledge, the only MR methods available to account for the correlation between SNPs are IVW, MR–Egger, and maximum likelihood-based MR, while there is a rapid progress in MR development method for uncorrelated SNP. This limits the use of other sensitivity tests such as MR Radial.

Therefore, close collaborations between methodologists, empirical researchers, and clinicians are required and will have the potential to improve the methodological issues and to ensure the strengths of MR findings.

Chapter 8 *Appendix*

Table A.1: MR results on effects of druggable genes on MS risk (nominally significant at p-value ≤ 0.05) (*part 1*)

Gene	Druggability tier	Outcome	No. SNP	Method	Beta	SE	p-value	FDR	Tissue
ACP2	3	MS risk (discovery)	3	IVW	-0.15	0.035	1.47E-05	3.72E-03	Blood
ACP2	3	MS risk (discovery)	3	MR-Egger	-0.224	0.136	1.00E-01		Blood
ALDH3A2	1	MS risk (discovery)	2	IVW	0.242	0.071	5.90E-04	4.14E-02	Blood
CCR4	1	MS risk (discovery)	1	Wald ratio	0.912	0.248	2.36E-04	2.39E-02	Blood
CCR4	1	MS risk (replication)	1	Wald ratio	0.966	0.251	1.15E-04	2.73E-04	Blood
CD6	1	MS risk (discovery)	13	IVW	-0.451	0.121	1.87E-04	1.98E-02	Blood
CD6	1	MS risk (discovery)	13	MR-Egger	-0.475	0.119	6.57E-05		Blood
CD6	1	MS risk (replication)	12	IVW	-0.353	0.075	2.73E-06	1.04E-05	Blood
CD6	1	MS risk (replication)	12	MR-Egger	-0.243	0.081	2.79E-03		Blood
CDK14	2	MS risk (discovery)	5	IVW	-0.757	0.188	5.88E-05	8.76E-03	Blood
CDK14	2	MS risk (discovery)	5	MR-Egger	-0.835	0.573	1.45E-01		Blood
CDK14	2	MS risk (replication)	2	IVW	0.425	0.215	4.79E-02	7.58E-02	Blood
CHEK2	1	MS risk (discovery)	1	Wald ratio	-1.274	0.284	7.11E-06	2.00E-03	Blood
CXCR4	1	MS risk (discovery)	1	Wald ratio	-1.815	0.439	3.55E-05	6.42E-03	Blood
DNMT3A	1	MS risk (discovery)	7	IVW	-0.759	0.158	1.48E-06	5.35E-04	Blood
DNMT3A	1	MS risk (discovery)	7	MR-Egger	-0.607	0.36	9.12E-02		Blood
GALK1	2	MS risk (discovery)	3	IVW	-0.247	0.071	5.43E-04	4.14E-02	Blood
GALK1	2	MS risk (discovery)	3	MR-Egger	-0.258	0.724	7.22E-01		Blood
HDAC3	1	MS risk (discovery)	1	Wald ratio	0.474	0.138	6.03E-04	4.14E-02	Blood
HSD3B7	4	MS risk (discovery)	2	IVW	0.535	0.141	1.41E-04	1.73E-02	Blood
IFITM1	3	MS risk (discovery)	3	IVW	0.295	0.078	1.56E-04	1.79E-02	Blood
IFITM1	3	MS risk (discovery)	3	MR-Egger	-0.318	17.744	9.86E-01		Blood
IFNGR2	1	MS risk (discovery)	11	IVW	0.111	0.031	3.31E-04	2.99E-02	Blood
IFNGR2	1	MS risk (discovery)	11	MR-Egger	0.061	0.05	2.27E-01		Blood
IFNGR2	1	MS risk (replication)	1	Wald ratio	0.147	0.037	7.43E-05	2.02E-04	Blood
IL7	3	MS risk (discovery)	1	Wald ratio	-1.508	0.268	1.94E-08	2.45E-05	Blood
IL7	3	MS risk (replication)	1	Wald ratio	-1.655	0.275	1.66E-09	1.05E-08	Blood
ITGA3	4	MS risk (discovery)	2	IVW	0.665	0.163	4.73E-05	7.99E-03	Blood
ITGB3	1	MS risk (discovery)	3	IVW	-0.362	0.093	9.47E-05	1.26E-02	Blood
ITGB3	1	MS risk (discovery)	3	MR-Egger	0.182	0.3	5.46E-01		Blood
JAK1	1	MS risk (discovery)	1	Wald ratio	-0.955	0.272	4.53E-04	3.83E-02	Blood
KLHL25	4	MS risk (discovery)	1	Wald ratio	1.125	0.301	1.87E-04	1.98E-02	Blood
KLHL3	4	MS risk (discovery)	11	IVW	-0.236	0.062	1.43E-04	1.73E-02	Blood
KLHL3	4	MS risk (discovery)	11	MR-Egger	-0.258	0.128	4.38E-02		Blood
MAP3K11	1	MS risk (discovery)	7	IVW	0.134	0.037	2.95E-04	2.77E-02	Blood
MAP3K11	1	MS risk (discovery)	7	MR-Egger	0.038	0.192	8.44E-01		Blood
MAP3K14	2	MS risk (discovery)	1	Wald ratio	1.223	0.232	1.31E-07	1.11E-04	Blood
MAPK11	1	MS risk (discovery)	1	Wald ratio	1.45	0.415	4.70E-04	3.84E-02	Blood
MAPK3	1	MS risk (discovery)	5	IVW	-0.178	0.035	2.79E-07	1.40E-04	Blood
MAPK3	1	MS risk (discovery)	5	MR-Egger	-0.154	0.14	2.71E-01		Blood
MAPK3	1	MS risk (replication)	2	IVW	-0.2	0.034	2.92E-09	1.39E-08	Blood
MAST3	3	MS risk (discovery)	3	IVW	0.431	0.084	3.32E-07	1.40E-04	Blood
MAST3	3	MS risk (discovery)	3	MR-Egger	0.986	0.356	5.58E-03		Blood
MAST3	3	MS risk (replication)	2	IVW	0.535	0.072	1.52E-13	1.45E-12	Blood

Table A.1: MR results on effects of druggable genes on MS risk (nominally significant at p-value ≤ 0.05) (*part 2*)

Gene	Druggability tier	Outcome	No. SNP	Method	Beta	SE	p-value	FDR	Tissue
MCL1	2	MS risk (discovery)	3	IVW	-0.342	0.094	2.92E-04	2.77E-02	Blood
MCL1	2	MS risk (discovery)	3	MR-Egger	-0.229	0.628	7.16E-01		Blood
MERTK	1	MS risk (discovery)	13	IVW	0.142	0.027	1.81E-07	1.15E-04	Blood
MERTK	1	MS risk (discovery)	13	MR-Egger	0.083	0.047	7.94E-02		Blood
NR1D1	2	MS risk (discovery)	3	IVW	0.338	0.079	2.10E-05	4.69E-03	Blood
NR1D1	2	MS risk (discovery)	3	MR-Egger	0.416	0.134	1.97E-03		Blood
PARP2	1	MS risk (discovery)	5	IVW	-0.19	0.055	6.06E-04	4.14E-02	Blood
PARP2	1	MS risk (discovery)	5	MR-Egger	-0.178	0.105	9.04E-02		Blood
PGLYRP1	4	MS risk (discovery)	6	IVW	-0.206	0.049	2.22E-05	4.69E-03	Blood
PGLYRP1	4	MS risk (discovery)	6	MR-Egger	-0.141	0.104	1.74E-01		Blood
SAE1	2	MS risk (discovery)	4	IVW	0.462	0.132	4.54E-04	3.83E-02	Blood
SAE1	2	MS risk (discovery)	4	MR-Egger	-0.044	0.293	8.81E-01		Blood
SIK3	1	MS risk (discovery)	2	IVW	1.678	0.418	5.83E-05	8.76E-03	Blood
SIK3	1	MS risk (replication)	1	Wald ratio	0.568	0.155	2.46E-04	4.25E-04	Blood
SLAMF7	1	MS risk (discovery)	3	IVW	0.456	0.096	1.88E-06	5.95E-04	Blood
SLAMF7	1	MS risk (discovery)	3	MR-Egger	0.365	0.512	4.76E-01		Blood
SLAMF7	1	MS risk (replication)	3	IVW	0.423	0.094	6.05E-06	1.92E-05	Blood
SLAMF7	1	MS risk (replication)	3	MR-Egger	0.39	0.58	5.01E-01		Blood
STAT3	1	MS risk (discovery)	4	IVW	-0.864	0.131	4.28E-11	1.08E-07	Blood
STAT3	1	MS risk (discovery)	4	MR-Egger	-1.133	0.985	2.50E-01		Blood
STAT3	1	MS risk (replication)	3	IVW	-0.492	0.131	1.66E-04	3.16E-04	Blood
STAT3	1	MS risk (replication)	3	MR-Egger	0.607	2.672	8.20E-01		Blood
TNFSF14	1	MS risk (discovery)	7	IVW	-0.51	0.121	2.71E-05	5.27E-03	Blood
TNFSF14	1	MS risk (discovery)	7	MR-Egger	-0.575	0.112	2.74E-07		Blood
TNFSF14	1	MS risk (replication)	2	IVW	-0.589	0.057	8.30E-25	1.58E-23	Blood
TYMP	2	MS risk (discovery)	10	IVW	-0.151	0.038	8.43E-05	1.18E-02	Blood
TYMP	2	MS risk (discovery)	10	MR-Egger	-0.11	0.065	8.95E-02		Blood
TYMP	2	MS risk (replication)	1	Wald ratio	-0.11	0.029	1.51E-04	3.16E-04	Blood
XPO1	2	MS risk (discovery)	1	Wald ratio	1.516	0.439	5.61E-04	4.14E-02	Blood
YARS	2	MS risk (discovery)	2	IVW	0.591	0.172	6.01E-04	4.14E-02	Blood
ABCA1	1	MS risk (discovery)	1	Wald ratio	-0.72	0.191	1.67E-04	2.29E-02	Brain
ALDH1L1	4	MS risk (discovery)	2	IVW	0.773	0.173	7.55E-06	1.41E-03	Brain
CDK3	2	MS risk (discovery)	1	Wald ratio	-0.539	0.147	2.36E-04	2.48E-02	Brain
DKKL1	3	MS risk (discovery)	1	Wald ratio	-0.764	0.12	2.25E-10	2.78E-07	Brain
GALC	3	MS risk (discovery)	1	Wald ratio	-0.391	0.107	2.41E-04	2.48E-02	Brain
GALC	3	MS risk (replication)	1	Wald ratio	-0.214	0.108	4.64E-02	4.64E-02	Brain
IFITM1	3	MS risk (discovery)	1	Wald ratio	1.195	0.264	5.91E-06	1.41E-03	Brain
IL2RA	1	MS risk (discovery)	1	Wald ratio	0.383	0.086	7.98E-06	1.41E-03	Brain
IL2RA	1	MS risk (replication)	1	Wald ratio	0.43	0.1	1.86E-05	2.79E-05	Brain
IL7	3	MS risk (discovery)	1	Wald ratio	-0.822	0.145	1.56E-08	6.41E-06	Brain
IL7	3	MS risk (replication)	1	Wald ratio	-0.896	0.171	1.62E-07	4.86E-07	Brain
KLHL18	4	MS risk (discovery)	2	IVW	0.37	0.1	2.25E-04	2.48E-02	Brain
MAPK3	1	MS risk (discovery)	1	Wald ratio	-0.819	0.174	2.44E-06	7.52E-04	Brain
PVR	3	MS risk (discovery)	1	Wald ratio	0.245	0.06	3.95E-05	6.09E-03	Brain
RPS6KB1	1	MS risk (discovery)	1	Wald ratio	-0.511	0.147	5.23E-04	4.97E-02	Brain

Table A.3: Heterogeneity and pleiotropy assessments for the effects of druggable genes on MS risk

Gene	Outcome	No.SNPs	MR-Egger intercept	MR-Egger intercept p-value	Cochrans' Q	Cochrans' Q p-value	I^2	Tissue
ACP2	MS risk (discovery)	3	0.035	5.46E-01	0.32	5.70E-01	0	Blood
CD6	MS risk (discovery)	13	0.003	7.68E-01	46.37	**2.78E-06**	0.74	Blood
CD6	MS risk (replication)	12	-0.029	**3.08E-02**	12.28	2.67E-01	0.1	Blood
CDK14	MS risk (discovery)	5	0.005	8.84E-01	3.88	2.75E-01	0	Blood
DNMT3A	MS risk (discovery)	7	-0.015	5.76E-01	2.21	8.19E-01	0	Blood
DNMT3A	MS risk (discovery)	7	-0.015	5.76E-01	2.21	8.19E-01	0	Blood
GALK1	MS risk (discovery)	3	0.003	9.86E-01	0.03	8.53E-01	0	Blood
IFITM1	MS risk (discovery)	3	0.166	9.70E-01	0.0001	9.92E-01	0	Blood
IFNGR2	MS risk (discovery)	12	0.027	1.85E-01	11.68	3.07E-01	0.06	Blood
ITGB3	MS risk (discovery)	3	-0.095	**4.42E-02**	4.1	**4.28E-02**	0.51	Blood
KLHL3	MS risk (discovery)	11	0.005	8.11E-01	4.33	8.89E-01	0	Blood
MAP3K11	MS risk (discovery)	7	0.03	6.07E-01	3.99	5.51E-01	0	Blood
MAPK3	MS risk (discovery)	5	-0.012	8.60E-01	1.68	6.42E-01	0	Blood
MAST3	MS risk (discovery)	3	-0.118	1.04E-01	2.1	1.47E-01	0.05	Blood
MCL1	MS risk (discovery)	3	-0.024	8.20E-01	0.06	8.04E-01	0	Blood
MERTK	MS risk (discovery)	13	0.026	8.97E-02	8.27	6.89E-01	0	Blood
NR1D1	MS risk (discovery)	3	-0.024	3.84E-01	0.63	4.27E-01	0	Blood
PARP2	MS risk (discovery)	5	-0.004	8.92E-01	2.79	4.24E-01	0	Blood
PGLYRP1	MS risk (discovery)	6	-0.024	3.98E-01	1.99	7.38E-01	0	Blood
SAE1	MS risk (discovery)	4	0.088	6.20E-02	1.16	5.60E-01	0	Blood
SLAMF7	MS risk (discovery)	3	0.015	8.38E-01	0.14	7.06E-01	0	Blood
SLAMF7	MS risk (replication)	3	0.005	9.48E-01	0.106	7.45E-01	0	Blood
STAT3	MS risk (discovery)	4	0.036	7.83E-01	3.3	1.92E-01	0.09	Blood
STAT3	MS risk (replication)	3	-0.153	6.78E-01	0.144	7.04E-01	0	Blood
TNFSF14	MS risk (discovery)	7	0.017	2.48E-01	13.37	**2.01E-02**	0.55	Blood
TYMP	MS risk (discovery)	10	-0.022	4.76E-01	21.18	**6.69E-03**	0.57	Blood

Table A.2 : MR results on effects of druggable genes on MS severity (nominally significant at p-value ≤ 0.05)

Gene	Druggability tier	Outcome	No. SNPs	Method	Beta	SE	p-value	FDR	Tissue
FFAR2	2	MS severity	1	Wald ratio	-2.361	0.637	2.09E-04	2.60E-01	Blood
MAPK3	1	MS severity	2	IVW	-0.26	0.071	2.42E-04	2.60E-01	Blood
MGAT2	2	MS severity	2	IVW	-1	0.283	4.11E-04	2.95E-01	Blood
RPS26	2	MS severity	2	IVW	1.469	0.431	6.48E-04	3.48E-01	Blood
PPID	1	MS severity	2	IVW	1.795	0.543	9.58E-04	4.12E-01	Blood
IFNGR1	1	MS severity	2	IVW	0.975	0.308	1.53E-03	5.49E-01	Blood
LTBR	1	MS severity	3	IVW	-0.54	0.176	2.13E-03	6.00E-01	Blood
RPS29	2	MS severity	1	Wald ratio	1.3	0.425	2.23E-03	6.00E-01	Blood
KLF5	2	MS severity	1	Wald ratio	-2.102	0.7	2.68E-03	6.41E-01	Blood
FAM3C	3	MS severity	1	Wald ratio	0.774	0.264	3.42E-03	6.90E-01	Blood
MYLK4	1	MS severity	9	IVW	0.273	0.094	3.76E-03	6.90E-01	Blood
MAPK13	1	MS severity	1	Wald ratio	1.667	0.577	3.85E-03	6.90E-01	Blood
SIGLEC9	2	MS severity	2	IVW	0.623	0.218	4.21E-03	6.97E-01	Blood
PLAT	1	MS severity	1	Wald ratio	-2.481	0.882	4.90E-03	7.53E-01	Blood
MAP3K8	2	MS severity	2	IVW	-0.861	0.31	5.40E-03	7.69E-01	Blood
PSMC3	1	MS severity	1	Wald ratio	-1.136	0.411	5.72E-03	7.69E-01	Blood
PAK2	1	MS severity	2	IVW	0.891	0.332	7.26E-03	7.81E-01	Blood
PDE6H	4	MS severity	2	IVW	1.136	0.425	7.55E-03	7.81E-01	Blood
TUBD1	4	MS severity	1	Wald ratio	-1.445	0.545	8.00E-03	7.81E-01	Blood
CTSS	1	MS severity	2	IVW	-0.348	0.133	8.81E-03	7.81E-01	Blood
NR3C2	1	MS severity	2	IVW	1.819	0.695	8.85E-03	7.81E-01	Blood
SIGLEC11	4	MS severity	2	IVW	-0.247	0.095	9.25E-03	7.81E-01	Blood
IL21R	3	MS severity	5	MR-Egger	-0.647	0.251	9.85E-03		Blood
CHSY1	3	MS severity	4	IVW	-0.236	0.092	1.01E-02	7.81E-01	Blood
LMTK3	4	MS severity	2	IVW	1.448	0.565	1.04E-02	7.81E-01	Blood
TNFRSF8	1	MS severity	1	Wald ratio	2.251	0.884	1.09E-02	7.81E-01	Blood
MAPK7	2	MS severity	2	IVW	1.156	0.454	1.09E-02	7.81E-01	Blood
LYPD5	4	MS severity	2	IVW	-0.836	0.332	1.18E-02	7.81E-01	Blood
CAC1D	#N/A	MS severity	1	Wald ratio	-1.539	0.612	1.20E-02	7.81E-01	Blood
FPR2	2	MS severity	5	IVW	-0.173	0.069	1.20E-02	7.81E-01	Blood
SPACA3	4	MS severity	3	IVW	0.796	0.317	1.21E-02	7.81E-01	Blood
VASH1	4	MS severity	3	IVW	-0.292	0.117	1.24E-02	7.81E-01	Blood
GMNN	1	MS severity	1	Wald ratio	-1.251	0.502	1.26E-02	7.81E-01	Blood
SPINK2	4	MS severity	3	IVW	-0.667	0.268	1.28E-02	7.81E-01	Blood
CTLA4	1	MS severity	1	Wald ratio	2.453	0.989	1.31E-02	7.81E-01	Blood
HIST1H3E	3	MS severity	2	IVW	0.287	0.116	1.33E-02	7.81E-01	Blood
FPR1	2	MS severity	11	IVW	-0.199	0.081	1.36E-02	7.81E-01	Blood
RPS2	2	MS severity	1	Wald ratio	1.467	0.598	1.42E-02	7.81E-01	Blood
ST6GAL1	3	MS severity	11	IVW	-0.426	0.174	1.44E-02	7.81E-01	Blood
ARSA	1	MS severity	2	IVW	-0.248	0.102	1.49E-02	7.81E-01	Blood
FCGR2A	3	MS severity	2	IVW	-1.457	0.602	1.54E-02	7.81E-01	Blood
MAP3K7	1	MS severity	3	MR-Egger	0.996	0.412	1.56E-02		Blood
CD34	3	MS severity	1	Wald ratio	1.264	0.524	1.58E-02	7.81E-01	Blood
CAMK1	2	MS severity	2	IVW	0.898	0.375	1.66E-02	7.81E-01	Blood

Table A.2 : MR results on effects of druggable genes on MS severity (nominally significant at p-value ≤ 0.05) *(continue)*

Gene	Druggability tier	Outcome	No. SNPs	Method	Beta	SE	p-value	FDR	Tissue
PEAK1	4	MS severity	3	IVW	-0.701	0.293	1.66E-02	7.81E-01	Blood
AEBP1	3	MS severity	1	Wald ratio	0.587	0.246	1.69E-02	7.81E-01	Blood
ABI3BP	4	MS severity	1	Wald ratio	1.311	0.551	1.74E-02	7.81E-01	Blood
PSMD4	1	MS severity	2	IVW	0.611	0.258	1.78E-02	7.81E-01	Blood
PPIF	2	MS severity	2	IVW	0.587	0.249	1.86E-02	7.81E-01	Blood
PRSS22	4	MS severity	2	IVW	0.455	0.193	1.86E-02	7.81E-01	Blood
APOL1	3	MS severity	2	IVW	-0.714	0.304	1.89E-02	7.81E-01	Blood
PVR	3	MS severity	1	Wald ratio	2.532	1.085	1.96E-02	7.81E-01	Blood
POLE2	1	MS severity	1	Wald ratio	-1.489	0.638	1.97E-02	7.81E-01	Blood
BCL2L1	1	MS severity	1	Wald ratio	1.952	0.838	1.98E-02	7.81E-01	Blood
CDK18	2	MS severity	1	Wald ratio	-2.137	0.921	2.03E-02	7.81E-01	Blood
BMP2K	1	MS severity	1	Wald ratio	-1.741	0.752	2.07E-02	1.03E-01	Brain
KLHL18	4	MS severity	1	Wald ratio	-1.237	0.537	2.12E-02	7.81E-01	Blood
VRK1	3	MS severity	3	IVW	0.452	0.196	2.14E-02	7.81E-01	Blood
KBTBD2	4	MS severity	1	Wald ratio	0.279	0.122	2.15E-02	7.81E-01	Blood
TIMP2	3	MS severity	4	IVW	-0.345	0.15	2.15E-02	7.81E-01	Blood
PRKCZ	1	MS severity	3	IVW	0.645	0.281	2.17E-02	7.81E-01	Blood
ICAM3	1	MS severity	2	IVW	0.298	0.13	2.18E-02	7.81E-01	Blood
EFEMP2	3	MS severity	2	IVW	-0.351	0.155	2.39E-02	8.32E-01	Blood
C1QB	4	MS severity	2	IVW	0.309	0.137	2.44E-02	8.32E-01	Blood
TXNRD1	1	MS severity	6	IVW	-0.498	0.223	2.51E-02	8.32E-01	Blood
RPS15A	2	MS severity	2	IVW	0.172	0.077	2.51E-02	8.32E-01	Blood
PLSCR3	#N/A	MS severity	2	IVW	0.288	0.129	2.54E-02	8.32E-01	Blood
PF4V1	4	MS severity	1	Wald ratio	-0.152	0.068	2.59E-02	8.32E-01	Blood
CYB5D2	4	MS severity	2	IVW	0.319	0.143	2.61E-02	8.32E-01	Blood
ME1	2	MS severity	1	Wald ratio	1.601	0.721	2.65E-02	8.32E-01	Blood
IKBKB	1	MS severity	1	Wald ratio	0.893	0.403	2.67E-02	8.32E-01	Blood
CPXM1	3	MS severity	1	Wald ratio	1.671	0.756	2.72E-02	8.35E-01	Blood
CDK14	2	MS severity	3	IVW	-1.333	0.607	2.79E-02	8.39E-01	Blood
PARP1	1	MS severity	2	IVW	0.257	0.117	2.83E-02	8.39E-01	Blood
SLAMF1	3	MS severity	4	IVW	-0.636	0.29	2.85E-02	8.39E-01	Blood
KDM3A	2	MS severity	2	IVW	1.052	0.485	3.02E-02	8.48E-01	Blood
IL21R	3	MS severity	5	IVW	-0.401	0.186	3.13E-02	8.48E-01	Blood
GDF9	4	MS severity	2	IVW	-0.893	0.416	3.18E-02	8.48E-01	Blood
BMP2K	1	MS severity	1	Wald ratio	-0.909	0.424	3.21E-02	8.48E-01	Blood
MAPKAPK5	2	MS severity	1	Wald ratio	-0.922	0.431	3.24E-02	8.48E-01	Blood
L3MBTL1	1	MS severity	1	Wald ratio	-2.051	0.962	3.31E-02	8.48E-01	Blood
FZD2	2	MS severity	2	IVW	-0.687	0.323	3.32E-02	8.48E-01	Blood
HLA-B	3	MS severity	2	IVW	0.25	0.118	3.39E-02	8.48E-01	Blood
RAF1	1	MS severity	3	IVW	0.674	0.32	3.49E-02	8.48E-01	Blood
PDE8A	1	MS severity	3	IVW	0.815	0.387	3.54E-02	8.48E-01	Blood
GHRL	1	MS severity	2	IVW	-0.489	0.232	3.54E-02	8.48E-01	Blood
PDE6B	4	MS severity	1	Wald ratio	0.576	0.274	3.57E-02	8.48E-01	Blood
TMEM9B	4	MS severity	2	IVW	-0.549	0.261	3.57E-02	8.48E-01	Blood

Table A.2 : MR results on effects of druggable genes on MS severity (nominally significant at p-value ≤ 0.05) *(continue)*

Gene	Druggability tier	Outcome	No. SNPs	Method	Beta	SE	p-value	FDR	Tissue
ALDH2	1	MS severity	1	Wald ratio	-1.745	0.831	3.58E-02	8.48E-01	Blood
RPS6KB1	1	MS severity	2	IVW	-0.781	0.373	3.61E-02	8.48E-01	Blood
TPP1	4	MS severity	2	IVW	0.536	0.256	3.63E-02	8.48E-01	Blood
NR3C1	1	MS severity	2	IVW	-1.327	0.635	3.65E-02	8.48E-01	Blood
FCRL5	3	MS severity	6	IVW	-0.161	0.077	3.68E-02	8.48E-01	Blood
IL15RA	3	MS severity	9	IVW	-0.284	0.136	3.70E-02	8.48E-01	Blood
HPSE	2	MS severity	3	IVW	-0.318	0.153	3.70E-02	8.48E-01	Blood
C1QC	4	MS severity	1	Wald ratio	1.005	0.482	3.70E-02	8.48E-01	Blood
ACVRL1	1	MS severity	3	IVW	-0.324	0.156	3.81E-02	8.63E-01	Blood
HSPB1	1	MS severity	1	Wald ratio	1.077	0.525	4.01E-02	8.63E-01	Blood
TNFSF10	2	MS severity	3	IVW	-0.487	0.238	4.05E-02	8.63E-01	Blood
FPGS	2	MS severity	1	Wald ratio	-1.743	0.851	4.07E-02	8.63E-01	Blood
CDK9	1	MS severity	1	Wald ratio	-1.779	0.869	4.07E-02	8.63E-01	Blood
TAOK3	1	MS severity	1	Wald ratio	-1.216	0.598	4.18E-02	8.63E-01	Blood
TNKS	2	MS severity	2	IVW	-1.165	0.572	4.18E-02	8.63E-01	Blood
HCAR3	2	MS severity	2	IVW	-0.174	0.086	4.25E-02	8.63E-01	Blood
BPNT1	4	MS severity	1	Wald ratio	1.375	0.678	4.27E-02	8.63E-01	Blood
IL17RB	4	MS severity	2	IVW	0.476	0.236	4.33E-02	8.63E-01	Blood
CD274	1	MS severity	2	IVW	0.441	0.219	4.35E-02	8.63E-01	Blood
ESR1	1	MS severity	3	IVW	-0.759	0.377	4.39E-02	8.63E-01	Blood
SLC5A11	2	MS severity	8	IVW	-0.236	0.118	4.43E-02	8.63E-01	Blood
SORD	2	MS severity	1	Wald ratio	1.334	0.664	4.47E-02	8.63E-01	Blood
ENTPD1	4	MS severity	6	IVW	0.198	0.099	4.53E-02	8.63E-01	Blood
TUBB6	1	MS severity	2	IVW	-0.157	0.078	4.53E-02	8.63E-01	Blood
ANGPT1	1	MS severity	10	IVW	0.426	0.214	4.61E-02	8.63E-01	Blood
CYP46A1	1	MS severity	1	Wald ratio	-0.823	0.413	4.62E-02	8.63E-01	Blood
FCRL6	3	MS severity	1	Wald ratio	-0.215	0.108	4.68E-02	8.63E-01	Blood
FCRL6	3	MS severity	1	Wald ratio	-0.215	0.108	4.68E-02	8.63E-01	Blood
TNFSF13	3	MS severity	2	IVW	0.258	0.13	4.68E-02	8.63E-01	Blood
COL4A3	3	MS severity	2	IVW	0.917	0.462	4.69E-02	8.63E-01	Blood
MAP3K7	1	MS severity	3	IVW	0.614	0.311	4.82E-02	8.63E-01	Blood
DHRS9	1	MS severity	3	IVW	-0.217	0.11	4.83E-02	8.63E-01	Blood
CASP6	2	MS severity	3	IVW	0.976	0.494	4.84E-02	8.63E-01	Blood
ALDH7A1	1	MS severity	5	IVW	0.308	0.156	4.86E-02	8.63E-01	Blood
CDK8	1	MS severity	2	IVW	-0.611	0.31	4.88E-02	8.63E-01	Blood
PDE5A	1	MS severity	3	IVW	-0.651	0.33	4.89E-02	8.63E-01	Blood
EPHB3	1	MS severity	1	Wald ratio	1.168	0.594	4.94E-02	8.63E-01	Blood
ANGPT1	1	MS severity	10	MR-Egger	0.579	0.317	6.81E-02		Blood
MYLK4	1	MS severity	9	MR-Egger	0.346	0.211	1.00E-01		Blood
ENTPD1	4	MS severity	6	MR-Egger	0.263	0.162	1.04E-01		Blood
TXNRD1	1	MS severity	6	MR-Egger	-0.493	0.304	1.05E-01		Blood
DHRS9	1	MS severity	3	MR-Egger	-0.512	0.32	1.10E-01		Blood
CHSY1	3	MS severity	4	MR-Egger	-0.225	0.151	1.36E-01		Blood

Table A.2 : MR results on effects of druggable genes on MS severity (nominally significant at p-value ≤ 0.05) *(continue)*

Gene	Druggability tier	Outcome	No. SNPs	Method	Beta	SE	p-value	FDR	Tissue
ACVRL1	1	MS severity	3	MR-Egger	-0.329	0.227	1.46E-01		Blood
GABRG2	1	MS severity	2	IVW	-0.719	0.512	1.60E-01	4.01E-01	Brain
FPR1	2	MS severity	11	MR-Egger	-0.169	0.131	1.97E-01		Blood
SLC5A11	2	MS severity	8	MR-Egger	-0.233	0.186	2.12E-01		Blood
HPSE	2	MS severity	3	MR-Egger	-0.358	0.366	3.28E-01		Blood
PRKCZ	1	MS severity	3	MR-Egger	2.124	2.221	3.39E-01		Blood
IL15RA	3	MS severity	9	MR-Egger	-0.424	0.483	3.80E-01		Blood
FAS	3	MS severity	1	Wald ratio	0.196	0.225	3.82E-01	6.37E-01	Brain
RAF1	1	MS severity	3	MR-Egger	1.082	1.254	3.88E-01		Blood
PDE5A	1	MS severity	3	MR-Egger	-0.67	0.897	4.55E-01		Blood
VRK1	3	MS severity	3	MR-Egger	0.691	0.988	4.84E-01		Blood
ST6GAL1	3	MS severity	11	MR-Egger	-0.355	0.652	5.86E-01		Blood
LTBR	1	MS severity	3	MR-Egger	-0.269	0.502	5.93E-01		Blood
ACKR2	4	MS severity	1	Wald ratio	0.16	0.305	6.00E-01	7.50E-01	Brain
TIMP2	3	MS severity	4	MR-Egger	-0.149	0.315	6.36E-01		Blood
FPR2	2	MS severity	5	MR-Egger	-0.158	0.335	6.38E-01		Blood
SLAMF1	3	MS severity	4	MR-Egger	-2.787	6.161	6.51E-01		Blood
SPINK2	4	MS severity	3	MR-Egger	-0.266	0.631	6.74E-01		Blood
CASP6	2	MS severity	3	MR-Egger	3.694	9.625	7.01E-01		Blood
ALDH7A1	1	MS severity	5	MR-Egger	0.144	0.411	7.26E-01		Blood
PDE8A	1	MS severity	3	MR-Egger	0.485	1.508	7.48E-01		Blood
PEAK1	4	MS severity	3	MR-Egger	-0.225	0.796	7.77E-01		Blood
TNFSF10	2	MS severity	3	MR-Egger	0.133	0.565	8.14E-01		Blood
CD109	4	MS severity	1	Wald ratio	-0.102	0.553	8.53E-01	8.53E-01	Brain
VASH1	4	MS severity	3	MR-Egger	-0.668	4.091	8.70E-01		Blood
FCRL5	3	MS severity	6	MR-Egger	-0.09	0.987	9.27E-01		Blood
ESR1	1	MS severity	3	MR-Egger	-0.958	13.777	9.45E-01		Blood
CDK14	2	MS severity	3	MR-Egger	-0.157	14.103	9.91E-01		Blood
SPACA3	4	MS severity	3	MR-Egger	0.384	481.093	9.99E-01		Blood

Table A.4: Heterogeneity and pleiotropy assessments for the effects of druggable genes on MS severity

Gene	Outcome	No.SNPs	MR-Egger intercept	MR-Egger intercept p-value	Cochrans' Q	Cochrans' Q p-value	i2	Tissue
ACVRL1	MS severity(discovery)	3	0.002	9.70E-01	1.94	1.64E-01	0	Blood
FPR1	MS severity(discovery)	11	-0.012	7.63E-01	8.68	4.68E-01	0	Blood
FPR2	MS severity(discovery)	5	-0.009	9.54E-01	0.39	9.41E-01	0	Blood
ALDH7A1	MS severity(discovery)	5	0.051	4.86E-01	0.64	8.87E-01	0	Blood
ANGPT1	MS severity(discovery)	10	-0.019	5.09E-01	8.55	3.82E-01	0	Blood
HPSE	MS severity(discovery)	3	0.011	9.00E-01	2.23	1.35E-01	0.1	Blood
IL15RA	MS severity(discovery)	9	0.031	7.42E-01	2.56	9.23E-01	0	Blood
IL21R	MS severity(discovery)	5	0.073	1.87E-01	3.4	3.34E-01	0	Blood
LTBR	MS severity(discovery)	3	-0.071	5.38E-01	0.52	4.69E-01	0	Blood
MAP3K7	MS severity(discovery)	3	-0.057	1.01E-01	1.43	2.32E-01	0	Blood
MYLK4	MS severity(discovery)	9	-0.03	6.13E-01	2.66	9.15E-01	0	Blood
PDE5A	MS severity(discovery)	3	0.003	9.81E-01	1.18	2.77E-01	0	Blood
PDE8A	MS severity(discovery)	3	0.043	8.10E-01	0.38	5.37E-01	0	Blood
PEAK1	MS severity(discovery)	3	-0.074	4.97E-01	0.59	4.42E-01	0	Blood
PRKCZ	MS severity(discovery)	3	-0.238	5.01E-01	0.77	3.81E-01	0	Blood
RAF1	MS severity(discovery)	3	-0.057	7.35E-01	1.46	2.27E-01	0	Blood
SLAMF1	MS severity(discovery)	4	0.27	7.22E-01	0.13	9.39E-01	0	Blood
SLC5A11	MS severity(discovery)	8	-0.005	8.41E-01	2.41	8.78E-01	0	Blood
SPACA3	MS severity(discovery)	3	0.045	9.99E-01	0	9.99E-01	0	Blood
SPINK2	MS severity(discovery)	3	-0.053	4.73E-01	1.43	2.32E-01	0	Blood
ST6GAL1	MS severity(discovery)	11	-0.012	8.99E-01	2.63	9.77E-01	0	Blood
TIMP2	MS severity(discovery)	4	-0.062	3.31E-01	0.64	7.24E-01	0	Blood
TNFSF10	MS severity(discovery)	3	-0.115	2.15E-01	1.5	2.20E-01	0	Blood
TXNRD1	MS severity(discovery)	6	-0.001	9.77E-01	3.83	4.29E-01	0	Blood
VASH1	MS severity(discovery)	3	0.112	9.26E-01	0.04	8.39E-01	0	Blood
VRK1	MS severity(discovery)	3	-0.059	7.88E-01	0.21	6.49E-01	0	Blood
CASP6	MS severity(discovery)	3	-0.198	7.73E-01	0.07	7.92E-01	0	Blood
CDK14	MS severity(discovery)	3	-0.077	9.32E-01	0.05	8.27E-01	0	Blood
CHSY1	MS severity(discovery)	4	-0.008	8.37E-01	0.84	6.57E-01	0	Blood
DHRS9	MS severity(discovery)	3	0.115	3.10E-01	2.21	1.37E-01	0.09	Blood
ENTPD1	MS severity(discovery)	6	-0.033	6.20E-01	6.56	1.61E-01	0.2	Blood
ESR1	MS severity(discovery)	3	0.02	9.87E-01	0	9.56E-01	0	Blood
FCRL5	MS severity(discovery)	6	-0.029	9.40E-01	0.23	9.94E-01	0	Blood

Bibliography

ACOSTA, J. N., SZEJKO, N. & FALCONE, G. J. 2021. Mendelian Randomization in Stroke: A Powerful Approach to Causal Inference and Drug Target Validation. *Frontiers in Genetics*, 1302.

AFIFI, S., MOHAMED, S., ZHAO, J. & FOSS, F. 2019. A drug safety evaluation of mogamulizumab for the treatment of cutaneous T-cell lymphoma. *Expert Opinion on Drug Safety,* 18, 769-776.

ALCINA, A., VANDENBROECK, K., OTAEGUI, D., SÁIZ, A., GONZALEZ, J. R., FERNANDEZ, O., CAVANILLAS, M., CÉNIT, M., ARROYO, R. & ALLOZA, I. 2010. The autoimmune disease-associated KIF5A, CD226 and SH2B3 gene variants confer susceptibility for multiple sclerosis. *Genes & Immunity,* 11, 439-445.

ALOIZOU, A.-M., SIOKAS, V., PATERAKI, G., LIAMPAS, I., BAKIRTZIS, C., TSOURIS, Z., LAZOPOULOS, G., CALINA, D., DOCEA, A. O. & TSATSAKIS, A. 2021. Thinking outside the ischemia box: advancements in the use of multiple sclerosis drugs in ischemic stroke. *Journal of Clinical Medicine,* 10, 630.

ANDERSEN, C. J., MURPHY, K. E. & FERNANDEZ, M. L. 2016. Impact of obesity and metabolic syndrome on immunity. *Advances in Nutrition,* 7, 66-75.

ANSELL, B. J., NAVAB, M., HAMA, S., KAMRANPOUR, N., FONAROW, G., HOUGH, G., RAHMANI, S., MOTTAHEDEH, R., DAVE, R. & REDDY, S. T. 2003. Inflammatory/antiinflammatory properties of high-density lipoprotein distinguish patients from control subjects better than high-density lipoprotein cholesterol levels and are favorably affected by simvastatin treatment. *Circulation,* 108, 2751-2756.

ANSELL, B. J., WATSON, K. E., FOGELMAN, A. M., NAVAB, M. & FONAROW, G. C. 2005. High-density lipoprotein function: recent advances. *Journal of the American College of Cardiology,* 46, 1792-1798.

AU YEUNG, S. L., BORGES, M.-C. & LAWLOR, D. A. 2018. Association of genetic instrumental variables for lung function on coronary artery disease risk: a 2-sample Mendelian randomization study. *Circulation: Genomic and Precision Medicine,* 11, e001952.

AZZARELLI, R., KERLOCH, T. & PACARY, E. 2015. Regulation of cerebral cortex development by Rho GTPases: insights from in vivo studies. *Frontiers in cellular neuroscience,* 8, 445.

BAECHER-ALLAN, C., KASKOW, B. J. & WEINER, H. L. 2018. Multiple sclerosis: mechanisms and immunotherapy. *Neuron,* 97, 742-768.

BAETTA, R., LENTO, S., GHILARDI, S., BARBATI, E., CORSINI, A., TREMOLI, E. & BANFI, C. 2015. Atorvastatin reduces long pentraxin 3 expression in vascular cells by inhibiting protein geranylgeranylation. *Vascular Pharmacology,* 67, 38-47.

BAIRD, D. A., LIU, J. Z., ZHENG, J., SIEBERTS, S. K., PERUMAL, T., ELSWORTH, B., RICHARDSON, T. G., CHEN, C.-Y., CARRASQUILLO, M. M. & ALLEN, M. 2021. Identifying drug targets for neurological and psychiatric disease via genetics and the brain transcriptome. *PLoS genetics,* 17, e1009224.

BARTER, P. J., BRANDRUP-WOGNSEN, G., PALMER, M. K. & NICHOLLS, S. J. 2010. Effect of statins on HDL-C: a complex process unrelated to changes in LDL-C: analysis of the VOYAGER Database. *Journal of lipid research,* 51, 1546-1553.

BAYO, J., FIORE, E. J., DOMINGUEZ, L. M., CANTERO, M. J., CIARLANTINI, M. S., MALVICINI, M., ATORRASAGASTI, C., GARCIA, M. G., ROSSI, M. & CAVASOTTO, C. 2021. Bioinformatic analysis of RHO family of GTPases identifies RAC1 pharmacological inhibition as a new therapeutic strategy for hepatocellular carcinoma. *Gut,* 70, 1362-1374.

BEECHAM, A. H., PATSOPOULOS, N. A., XIFARA, D. K., DAVIS, M. F., KEMPPINEN, A., COTSAPAS, C., SHAH, T. S., SPENCER, C., BOOTH, D. & GORIS, A. 2013a. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature genetics,* 45, 1353.

BEECHAM, A. H., PATSOPOULOS, N. A., XIFARA, D. K., DAVIS, M. F., KEMPPINEN, A., COTSAPAS, C., SHAHI, T. S., SPENCER, C., BOOTH, D. & GORIS, A. 2013b. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature genetics,* 45, 1353.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological),* 57**,** 289-300.
BENN, M. & NORDESTGAARD, B. G. 2018. From genome-wide association studies to Mendelian randomization: novel opportunities for understanding cardiovascular disease causality, pathogenesis, prevention, and treatment. *Cardiovascular research,* 114**,** 1192-1208.
BENN, M., NORDESTGAARD, B. G., FRIKKE-SCHMIDT, R. & TYBJÆRG-HANSEN, A. 2017. Low LDL cholesterol, PCSK9 and HMGCR genetic variation, and risk of Alzheimer's disease and Parkinson's disease: Mendelian randomisation study. *bmj,* 357.
BENNETT, D. A. & HOLMES, M. V. 2017. Mendelian randomisation in cardiovascular research: an introduction for clinicians. *Heart***,** heartjnl-2016-310605.
BJORNEVIK, K., CORTESE, M., HEALY, B. C., KUHLE, J., MINA, M. J., LENG, Y., ELLEDGE, S. J., NIEBUHR, D. W., SCHER, A. I. & MUNGER, K. L. 2022. Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science,* 375**,** 296-301.
BOOT, E., EKKER, M. S., PUTAALA, J., KITTNER, S., DE LEEUW, F.-E. & TULADHAR, A. M. 2020. Ischaemic stroke in young adults: a global perspective. *Journal of Neurology, Neurosurgery & Psychiatry,* 91**,** 411-417.
BOWDEN, J., DAVEY SMITH, G. & BURGESS, S. 2015. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International journal of epidemiology,* 44**,** 512-525.
BOWDEN, J., DAVEY SMITH, G., HAYCOCK, P. C. & BURGESS, S. 2016. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genetic epidemiology,* 40**,** 304-314.
BOWDEN, J., DEL GRECO M, F., MINELLI, C., DAVEY SMITH, G., SHEEHAN, N. & THOMPSON, J. 2017. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Statistics in medicine,* 36**,** 1783-1802.
BOWDEN, J., HEMANI, G. & DAVEY SMITH, G. 2018a. Invited commentary: detecting individual and global horizontal pleiotropy in Mendelian randomization—a job for the humble heterogeneity statistic? *American journal of epidemiology,* 187**,** 2681-2685.
BOWDEN, J. & HOLMES, M. V. 2019. Meta-analysis and Mendelian randomization: A review. *Research synthesis methods,* 10**,** 486-496.
BOWDEN, J., SPILLER, W., DEL GRECO M, F., SHEEHAN, N., THOMPSON, J., MINELLI, C. & DAVEY SMITH, G. 2018b. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *International journal of epidemiology,* 47**,** 1264-1278.
BRIGGS, F., SHAO, X., GOLDSTEIN, B. A., OKSENBERG, J. R., BARCELLOS, L. F. & DE JAGER, P. L. 2011. Genome-wide association study of severity in multiple sclerosis. *Genes & Immunity,* 12.
BROWNLEE, W. J., HARDY, T. A., FAZEKAS, F. & MILLER, D. H. 2017. Diagnosis of multiple sclerosis: progress and challenges. *The Lancet,* 389**,** 1336-1346.
BUECHLER, C., KRAUTBAUER, S. & EISINGER, K. 2015. Adipose tissue fibrosis. *World journal of diabetes,* 6**,** 548.
BURGESS, S. & BOWDEN, J. 2015. Integrating summarized data from multiple genetic variants in Mendelian randomization: bias and coverage properties of inverse-variance weighted methods. *arXiv preprint arXiv:1512.04486*.
BURGESS, S., BOWDEN, J., FALL, T., INGELSSON, E. & THOMPSON, S. G. 2017a. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology (Cambridge, Mass.),* 28**,** 30.
BURGESS, S., BUTTERWORTH, A. & THOMPSON, S. G. 2013. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology,* 37**,** 658-665.
BURGESS, S., DUDBRIDGE, F. & THOMPSON, S. G. 2016. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Statistics in medicine,* 35**,** 1880-1906.
BURGESS, S., SCOTT, R. A., TIMPSON, N. J., SMITH, G. D., THOMPSON, S. G. & CONSORTIUM, E.-I. 2015. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *European journal of epidemiology,* 30**,** 543-552.

BURGESS, S., SMALL, D. S. & THOMPSON, S. G. 2017b. A review of instrumental variable estimators for Mendelian randomization. *Statistical methods in medical research,* 26**,** 2333-2355.
BURGESS, S., SMITH, G. D., DAVIES, N. M., DUDBRIDGE, F., GILL, D., GLYMOUR, M. M., HARTWIG, F. P., HOLMES, M. V., MINELLI, C. & RELTON, C. L. 2019. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Research,* 4.
BURGESS, S. & THOMPSON, S. G. 2015. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *American journal of epidemiology,* 181**,** 251-260.
BURGESS, S. & THOMPSON, S. G. 2017. Interpreting findings from Mendelian randomization using the MR-Egger method. *European journal of epidemiology,* 32**,** 377-389.
BURGESS, S., THOMPSON, S. G. & COLLABORATION, C. C. G. 2011. Avoiding bias from weak instruments in Mendelian randomization studies. *International journal of epidemiology,* 40**,** 755-764.
BURGESS, S., ZUBER, V., VALDES-MARQUEZ, E., SUN, B. B. & HOPEWELL, J. C. 2017c. Mendelian randomization with fine-mapped genetic data: choosing from large numbers of correlated instrumental variables. *Genetic epidemiology,* 41**,** 714-725.
CAI, A., ZHOU, Y. & LI, L. 2015. Rho-GTPase and atherosclerosis: pleiotropic effects of statins. *Journal of the American Heart Association,* 4**,** e002113.
CAPRIO, M., RUSSO, C., GIUGLIANO, A., RAGUCCI, M. & MANCINI, M. 2016. Vascular disease in patients with multiple sclerosis: a review. *J Vasc Med Surg,* 4**,** 746-753.
CARRERAS-TORRES, R., JOHANSSON, M., HAYCOCK, P. C., RELTON, C. L., SMITH, G. D., BRENNAN, P. & MARTIN, R. M. 2018. Role of obesity in smoking behaviour: Mendelian randomisation study in UK Biobank. *Bmj,* 361.
CARTER, P., VITHAYATHIL, M., KAR, S., POTLURI, R., MASON, A. M., LARSSON, S. C. & BURGESS, S. 2020. Predicting the effect of statins on cancer risk using genetic variants from a Mendelian randomization study in the UK Biobank. *Elife,* 9**,** e57191.
CELARAIN, N. & TOMAS-ROIG, J. 2020. Aberrant DNA methylation profile exacerbates inflammation and neurodegeneration in multiple sclerosis patients. *Journal of neuroinflammation,* 17**,** 1-17.
CENERI, N., ZHAO, L., YOUNG, B. D., HEALY, A., COSKUN, S., VASAVADA, H., YAROVINSKY, T. O., IKE, K., PARDI, R. & QIN, L. 2017. Rac2 modulates atherosclerotic calcification by regulating macrophage interleukin-1β production. *Arteriosclerosis, thrombosis, and vascular biology,* 37**,** 328-340.
CERVANTES-GRACIA, K. & HUSI, H. 2018. Integrative analysis of multiple sclerosis using a systems biology approach. *Scientific reports,* 8**,** 1-14.
CHA, Y., EREZ, T., REYNOLDS, I., KUMAR, D., ROSS, J., KOYTIGER, G., KUSKO, R., ZESKIND, B., RISSO, S. & KAGAN, E. 2018. Drug repurposing from the perspective of pharmaceutical companies. *British journal of pharmacology,* 175**,** 168-180.
CHANDRA, A. & XU, Y. M. 2016. Cholesterol: A necessary evil from a multiple sclerosis perspective. *Clinical and Experimental Neuroimmunology,* 7**,** 145-157.
CHATAWAY, J., SCHUERER, N., ALSANOUSI, A., CHAN, D., MACMANUS, D., HUNTER, K., ANDERSON, V., BANGHAM, C. R., CLEGG, S. & NIELSEN, C. 2014. Effect of high-dose simvastatin on brain atrophy and disability in secondary progressive multiple sclerosis (MS-STAT): a randomised, placebo-controlled, phase 2 trial. *The Lancet,* 383**,** 2213-2221.
CONSORTIUM, G. 2020. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science,* 369**,** 1318-1330.
CONSORTIUM*†, I. M. S. G., ANZGENE, IIBDGC & WTCCC2 2019. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science,* 365**,** eaav7188.
CONWAY, D. S., THOMPSON, N. R. & COHEN, J. A. 2017. Influence of hypertension, diabetes, hyperlipidemia, and obstructive lung disease on multiple sclerosis disease course. *Multiple Sclerosis Journal,* 23**,** 277-285.
CORDLE, A., KOENIGSKNECHT-TALBOO, J., WILKINSON, B., LIMPERT, A. & LANDRETH, G. 2005. Mechanisms of statin-mediated inhibition of small G-protein function. *Journal of Biological Chemistry,* 280**,** 34202-34209.
CORDLE, A. & LANDRETH, G. 2005. 3-Hydroxy-3-methylglutaryl-coenzyme A reductase inhibitors attenuate β-amyloid-induced microglial inflammatory responses. *Journal of Neuroscience,* 25**,** 299-307.
DARLING, N. J., TOTH, R., ARTHUR, J. S. C. & CLARK, K. 2017. Inhibition of SIK2 and SIK3 during differentiation enhances the anti-inflammatory phenotype of macrophages. *Biochemical Journal,* 474**,** 521-537.

DAVEY SMITH, G. & HEMANI, G. 2014. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human molecular genetics,* 23**,** R89-R98.

DE ANGELIS, F., JOHN, N. A. & BROWNLEE, W. J. 2018. Disease-modifying therapies for multiple sclerosis. *bmj,* 363.

DE HEREDIA, F. P., GÓMEZ-MARTÍNEZ, S. & MARCOS, A. 2012. Obesity, inflammation and the immune system. *Proceedings of the Nutrition Society,* 71**,** 332-338.

DE SILVA, N. M. G., FREATHY, R. M., PALMER, T. M., DONNELLY, L. A., LUAN, J. A., GAUNT, T., LANGENBERG, C., WEEDON, M. N., SHIELDS, B. & KNIGHT, B. A. 2011. Mendelian randomization studies do not support a role for raised circulating triglyceride levels influencing type 2 diabetes, glucose levels, or insulin resistance. *Diabetes,* 60**,** 1008-1018.

DELBOSC, S., MORENA, M., DJOUAD, F., LEDOUCEN, C., DESCOMPS, B. & CRISTOL, J.-P. 2002. Statins, 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors, are able to reduce superoxide anion production by NADPH oxidase in THP-1-derived monocytes. *Journal of cardiovascular pharmacology,* 40**,** 611-617.

DEVORAK, J., MOKRY, L. E., MORRIS, J. A., FORGETTA, V., DAVEY SMITH, G., SAWCER, S. & RICHARDS, J. B. 2017. Large differences in adiponectin levels have no clear effect on multiple sclerosis risk: A Mendelian randomization study. *Multiple Sclerosis Journal,* 23**,** 1461-1468.

DIDONNA, A. & OKSENBERG, J. R. 2015. Genetic determinants of risk and progression in multiple sclerosis. *Clinica chimica acta,* 449**,** 16-22.

DURANT, R., KLOUCHE, K., DELBOSC, S., MORENA, M., AMIGUES, L., BERAUD, J. J., CANAUD, B. & CRISTOL, J. P. 2004. Superoxide anion overproduction in sepsis: effects of vitamin E and simvastatin. *Shock,* 22**,** 34-39.

DURINCK, S., MOREAU, Y., KASPRZYK, A., DAVIS, S., DE MOOR, B., BRAZMA, A. & HUBER, W. 2005. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics,* 21**,** 3439-3440.

ENDRES, M., LAUFS, U., HUANG, Z., NAKAMURA, T., HUANG, P., MOSKOWITZ, M. A. & LIAO, J. K. 1998. Stroke protection by 3-hydroxy-3-methylglutaryl (HMG)-CoA reductase inhibitors mediated by endothelial nitric oxide synthase. *Proceedings of the National Academy of Sciences,* 95**,** 8880-8885.

ESHAGHI, A., KIEVIT, R. A., PRADOS, F., SUDRE, C. H., NICHOLAS, J., CARDOSO, M. J., CHAN, D., NICHOLAS, R., OURSELIN, S. & GREENWOOD, J. 2019. Applying causal models to explore the mechanism of action of simvastatin in progressive multiple sclerosis. *Proceedings of the National Academy of Sciences,* 116**,** 11020-11027.

FATTOUH, R., GUO, C.-H., LAM, G. Y., GAREAU, M. G., NGAN, B.-Y., GLOGAUER, M., MUISE, A. M. & BRUMELL, J. H. 2013. Rac2-deficiency leads to exacerbated and protracted colitis in response to Citrobacter rodentium infection. *PloS one,* 8**,** e61629.

FERENCE, B. A., MAJEED, F., PENUMETCHA, R., FLACK, J. M. & BROOK, R. D. 2015. Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or both: a 2× 2 factorial Mendelian randomization study. *Journal of the American College of Cardiology,* 65**,** 1552-1561.

FERENCE, B. A., YOO, W., ALESH, I., MAHAJAN, N., MIROWSKA, K. K., MEWADA, A., KAHN, J., AFONSO, L., WILLIAMS, K. A. & FLACK, J. M. 2012. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. *Journal of the American College of Cardiology,* 60**,** 2631-2639.

FINAN, C., GAULTON, A., KRUGER, F. A., LUMBERS, R. T., SHAH, T., ENGMANN, J., GALVER, L., KELLEY, R., KARLSSON, A. & SANTOS, R. 2017. The druggable genome and support for target identification and validation in drug development. *Science translational medicine,* 9**,** eaag1166.

FLISTER, M. J., HOFFMAN, M. J., LEMKE, A., PRISCO, S. Z., RUDEMILLER, N., O'MEARA, C. C., TSAIH, S.-W., MORENO, C., GEURTS, A. M. & LAZAR, J. 2015. SH2B3 is a genetic determinant of cardiac inflammation and fibrosis. *Circulation: Cardiovascular Genetics,* 8**,** 294-304.

FOLEY, C. N., MASON, A. M., KIRK, P. D. & BURGESS, S. 2021. MR-Clust: clustering of genetic variants in Mendelian randomization with similar causal estimates. *Bioinformatics,* 37**,** 531-541.

FORD, H. 2020. Clinical presentation and diagnosis of multiple sclerosis. *Clinical Medicine,* 20**,** 380.

FORDE, E. A., DOGAN, R.-N. E. & KARPUS, W. J. 2011. CCR4 contributes to the pathogenesis of experimental autoimmune encephalomyelitis by regulating inflammatory macrophage function. *Journal of neuroimmunology,* 236**,** 17-26.

FORNY-GERMANO, L., DE FELICE, F. G. & VIEIRA, M. N. D. N. 2019. The role of leptin and adiponectin in obesity-associated cognitive decline and Alzheimer's disease. *Frontiers in neuroscience***,** 1027.

FORSYTHE, L. K., WALLACE, J. M. & LIVINGSTONE, M. B. E. 2008. Obesity and inflammation: the effects of weight loss. *Nutrition research reviews,* 21**,** 117-133.

FRANKENFIELD, D. C., ROWE, W. A., COONEY, R. N., SMITH, J. S. & BECKER, D. 2001. Limits of body mass index to detect obesity and predict body composition. *Nutrition,* 17**,** 26-30.

FUJINAMI, R. S., VON HERRATH, M. G., CHRISTEN, U. & WHITTON, J. L. 2006. Molecular mimicry, bystander activation, or viral persistence: infections and autoimmune disease. *Clinical microbiology reviews,* 19**,** 80-94.

GAFSON, A., THORNE, T., MCKECHNIE, C., JIMENEZ, B., NICHOLAS, R. & MATTHEWS, P. 2018. Lipoprotein markers associated with disability from multiple sclerosis. *Scientific reports,* 8**,** 1-8.

GAJOFATTO, A. & BENEDETTI, M. D. 2015. Treatment strategies for multiple sclerosis: when to start, when to change, when to stop? *World Journal of Clinical Cases: WJCC,* 3**,** 545.

GAZIANO, L., GIAMBARTOLOMEI, C., PEREIRA, A. C., GAULTON, A., POSNER, D. C., SWANSON, S. A., HO, Y.-L., IYENGAR, S. K., KOSIK, N. M. & VUJKOVIC, M. 2021. Actionable druggable genome-wide Mendelian randomization identifies repurposing opportunities for COVID-19. *Nature medicine,* 27**,** 668-676.

GEORGAKIS, M. K., GILL, D., WEBB, A. J., EVANGELOU, E., ELLIOTT, P., SUDLOW, C. L., DEHGHAN, A., MALIK, R., TZOULAKI, I. & DICHGANS, M. 2020. Genetically determined blood pressure, antihypertensive drug classes, and risk of stroke subtypes. *Neurology,* 95**,** e353-e361.

GHASEMI, N., RAZAVI, S. & NIKZAD, E. 2017. Multiple sclerosis: pathogenesis, symptoms, diagnoses and cell-based therapy. *Cell Journal (Yakhteh),* 19**,** 1.

GILGUN-SHERKI, Y., MELAMED, E. & OFFEN, D. 2004. The role of oxidative stress in the pathogenesis of multiple sclerosis: the need for effective antioxidant therapy. *Journal of neurology,* 251**,** 261-268.

GILL, D., GEORGAKIS, M. K., KOSKERIDIS, F., JIANG, L., FENG, Q., WEI, W.-Q., THEODORATOU, E., ELLIOTT, P., DENNY, J. C. & MALIK, R. 2019. Use of genetic variants related to antihypertensive drugs to inform on efficacy and side effects. *Circulation,* 140**,** 270-279.

GILL, D., GEORGAKIS, M. K., WALKER, V. M., SCHMIDT, A. F., GKATZIONIS, A., FREITAG, D. F., FINAN, C., HINGORANI, A. D., HOWSON, J. M. & BURGESS, S. 2021. Mendelian randomization for studying the effects of perturbing drug targets. *Wellcome open research,* 6.

GRECO, M., DEL, F., MINELLI, C., SHEEHAN, N. A. & THOMPSON, J. R. 2015. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Statistics in medicine,* 34**,** 2926-2940.

GREENWOOD, J., STEINMAN, L. & ZAMVIL, S. S. 2006. Statin therapy and autoimmune disease: from protein prenylation to immunomodulation. *Nature Reviews Immunology,* 6**,** 358.

GREGORY, S. G., SCHMIDT, S., SETH, P., OKSENBERG, J. R., HART, J., PROKOP, A., CAILLIER, S. J., BAN, M., GORIS, A. & BARCELLOS, L. F. 2007. Interleukin 7 receptor α chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nature genetics,* 39**,** 1083-1091.

GREMESE, E., TOLUSSO, B., GIGANTE, M. R. & FERRACCIOLI, G. 2014. Obesity as a risk and severity factor in rheumatic diseases (autoimmune chronic inflammatory diseases). *Frontiers in immunology,* 5**,** 576.

GRIFFITH, G. J., MORRIS, T. T., TUDBALL, M. J., HERBERT, A., MANCANO, G., PIKE, L., SHARP, G. C., STERNE, J., PALMER, T. M. & DAVEY SMITH, G. 2020. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nature communications,* 11**,** 1-12.

GROENWOLD, R. H., GOEMAN, J. J., LE CESSIE, S. & DEKKERS, O. M. 2021. Multiple testing: when is many too much? *European journal of endocrinology,* 184**,** E11-E14.

HAASE, C. L., LOPES, S., OLSEN, A. H., SATYLGANOVA, A., SCHNECKE, V. & MCEWAN, P. 2021. Weight loss and risk reduction of obesity-related outcomes in 0.5 million people: evidence from a UK primary care database. *International Journal of Obesity,* 45**,** 1249-1258.

HABIB, A., SHAMSEDDEEN, I., NASRALLAH, M. S., ANTOUN, T. A., NEMER, G., BERTOGLIO, J., BADREDDINE, R. & BADR, K. F. 2007. Modulation of COX-2 expression by statins in human monocytic cells. *The FASEB Journal,* 21**,** 1665-1674.

HARBO, H. F., GOLD, R. & TINTORÉ, M. 2013. Sex and gender issues in multiple sclerosis. *Therapeutic advances in neurological disorders,* 6**,** 237-248.
HARROUD, A., MITCHELL, R. E., RICHARDSON, T. G., MORRIS, J. A., FORGETTA, V., DAVEY SMITH, G., BARANZINI, S. E. & RICHARDS, J. B. 2021. Childhood obesity and multiple sclerosis: A Mendelian randomization study. *Multiple Sclerosis Journal***,** 13524585211001781.
HARROUD, A., MORRIS, J. A., FORGETTA, V., MITCHELL, R., SMITH, G. D., SAWCER, S. J. & RICHARDS, J. B. 2019. Effect of age at puberty on risk of multiple sclerosis: A mendelian randomization study. *Neurology,* 92**,** e1803-e1810.
HARTUNG, D. M. 2021. Health economics of disease-modifying therapy for multiple sclerosis in the United States. *Therapeutic Advances in Neurological Disorders,* 14**,** 1756286420987031.
HAUSER, S. L. & CREE, B. A. 2020. Treatment of multiple sclerosis: a review. *The American Journal of Medicine*.
HEDSTRÖM, A. K., OLSSON, T. & ALFREDSSON, L. 2012. High body mass index before age 20 is associated with increased risk for multiple sclerosis in both men and women. *Multiple Sclerosis Journal,* 18**,** 1334-1336.
HEDSTRÖM, A. K., SUNDQVIST, E., BÄÄRNHIELM, M., NORDIN, N., HILLERT, J., KOCKUM, I., OLSSON, T. & ALFREDSSON, L. 2011. Smoking and two human leukocyte antigen genes interact to increase the risk for multiple sclerosis. *Brain,* 134**,** 653-664.
HEMANI, G., BOWDEN, J. & DAVEY SMITH, G. 2018a. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Human molecular genetics,* 27**,** R195-R208.
HEMANI, G., ZHENG, J., ELSWORTH, B., WADE, K. H., HABERLAND, V., BAIRD, D., LAURIN, C., BURGESS, S., BOWDEN, J. & LANGDON, R. 2018b. The MR-Base platform supports systematic causal inference across the human phenome. *eLife,* 7**,** e34408.
HIGGINS, J. P., THOMPSON, S. G., DEEKS, J. J. & ALTMAN, D. G. 2003. Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal,* 327**,** 557.
HINDY, G., ENGSTRÖM, G., LARSSON, S. C., TRAYLOR, M., MARKUS, H. S., MELANDER, O. & ORHO-MELANDER, M. 2018. Role of blood lipids in the development of ischemic stroke and its subtypes: a Mendelian randomization study. *Stroke,* 49**,** 820-827.
HODGE, R. G. & RIDLEY, A. J. 2016. Regulating Rho GTPases and their regulators. *Nature reviews Molecular cell biology,* 17**,** 496.
HØGLUND, R. A. & MAGHAZACHI, A. A. 2014. Multiple sclerosis and the role of immune cells. *World journal of experimental medicine,* 4**,** 27.
HONG, Y., TANG, H. R., MA, M., CHEN, N., XIE, X. & HE, L. 2019. Multiple sclerosis and stroke: a systematic review and meta-analysis. *BMC neurology,* 19**,** 1-11.
HOWELL, A. E., ZHENG, J., HAYCOCK, P. C., MCALEENAN, A., RELTON, C., MARTIN, R. M. & KURIAN, K. M. 2018. Use of Mendelian randomization for identifying risk factors for brain tumors. *Frontiers in genetics,* 9**,** 525.
HOWRIGAN, L. A. D. 2019. GWAS OF UK BIOBANK BIOMARKER MEASUREMENTS.
HUAN, T., JOEHANES, R., SONG, C., PENG, F., GUO, Y., MENDELSON, M., YAO, C., LIU, C., MA, J. & RICHARD, M. 2019. Genome-wide identification of DNA methylation QTLs in whole blood highlights pathways for cardiovascular disease. *Nature communications,* 10**,** 1-14.
IADECOLA, C. & ANRATHER, J. 2011. The immunology of stroke: from mechanisms to translation. *Nature medicine,* 17**,** 796-808.
JACOBS, B. M., NOYCE, A. J., BESTWICK, J., BELETE, D., GIOVANNONI, G. & DOBSON, R. 2021. Gene-Environment Interactions in Multiple Sclerosis: A UK Biobank Study. *Neurology-Neuroimmunology Neuroinflammation,* 8.
JACOBS, B. M., TAYLOR, T., AWAD, A., BAKER, D., GIOVANONNI, G., NOYCE, A. J. & DOBSON, R. 2020. Summary-data-based Mendelian randomization prioritizes potential druggable targets for multiple sclerosis. *Brain communications,* 2**,** fcaa119.
JAUCH, E. C., LINDSELL, C., BRODERICK, J., FAGAN, S. C., TILLEY, B. C. & LEVINE, S. R. 2006. Association of serial biochemical markers with acute ischemic stroke: the National Institute of Neurological Disorders and Stroke recombinant tissue plasminogen activator Stroke Study. *Stroke,* 37**,** 2508-2513.
JAVIDI, E. & MAGNUS, T. 2019. Autoimmunity after ischemic stroke and brain injury. *Frontiers in immunology,* 10**,** 686.

JIANG, X., OLSSON, T. & ALFREDSSON, L. 2018. Age at menarche and risk of multiple sclerosis: current progress from epidemiological investigations. *Frontiers in immunology,* 9**,** 2600.

JIN, H. Y., TUDOR, Y., CHOI, K., SHAO, Z., SPARLING, B. A., MCGIVERN, J. G. & SYMONS, A. 2020. High-Throughput Implementation of the NanoBRET Target Engagement Intracellular Kinase Assay to Reveal Differential Compound Engagement by SIK2/3 Isoforms. *SLAS DISCOVERY: Advancing the Science of Drug Discovery,* 25**,** 215-222.

JONNALAGADDA, S. S., SKINNER, R. & MOORE, L. 2004. Overweight athlete: fact or fiction? *Current sports medicine reports,* 3**,** 198-205.

JORISSEN, W., WOUTERS, E., BOGIE, J. F., VANMIERLO, T., NOBEN, J.-P., SVIRIDOV, D., HELLINGS, N., SOMERS, V., VALCKE, R. & VANWIJMEERSCH, B. 2017. Relapsing-remitting multiple sclerosis patients display an altered lipoprotein profile with dysfunctional HDL. *Scientific reports,* 7**,** 1-14.

KANNEGANTI, T.-D. & DIXIT, V. D. 2012. Immunological complications of obesity. *Nature immunology,* 13**,** 707-712.

KAPPOS, L., BAR-OR, A., CREE, B. A., FOX, R. J., GIOVANNONI, G., GOLD, R., VERMERSCH, P., ARNOLD, D. L., ARNOULD, S. & SCHERZ, T. 2018. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. *The Lancet,* 391**,** 1263-1273.

KING, E. A., DAVIS, J. W. & DEGNER, J. F. 2019. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS genetics,* 15**,** e1008489.

KISTER, I. & KANTARCI, O. H. 2020. Multiple Sclerosis Severity Score: Concept and applications. *Multiple Sclerosis Journal,* 26**,** 548-553.

KLINEOVA, S. & LUBLIN, F. D. 2018. Clinical course of multiple sclerosis. *Cold Spring Harbor perspectives in medicine,* 8**,** a028928.

KLINGBERG, E., BILBERG, A., BJÖRKMAN, S., HEDBERG, M., JACOBSSON, L., FORSBLAD-D'ELIA, H., CARLSTEN, H., ELIASSON, B. & LARSSON, I. 2019. Weight loss improves disease activity in patients with psoriatic arthritis and obesity: an interventional study. *Arthritis research & therapy,* 21**,** 1-10.

KORTHAUER, K., KIMES, P. K., DUVALLET, C., REYES, A., SUBRAMANIAN, A., TENG, M., SHUKLA, C., ALM, E. J. & HICKS, S. C. 2019. A practical guide to methods controlling false discoveries in computational biology. *Genome biology,* 20**,** 1-21.

KOSMAS, C. E., MARTINEZ, I., SOURLAS, A., BOUZA, K. V., CAMPOS, F. N., TORRES, V., MONTAN, P. D. & GUZMAN, E. 2018. High-density lipoprotein (HDL) functionality and its relevance to atherosclerotic cardiovascular disease. *Drugs in context,* 7.

KURODA, M. & SAKAUE, H. 2017. Adipocyte death and chronic inflammation in obesity. *The journal of medical investigation,* 64**,** 193-196.

KURTZKE, J. F. 1983. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology,* 33**,** 1444-1444.

KUSMINSKI, C. M., BICKEL, P. E. & SCHERER, P. E. 2016. Targeting adipose tissue in the treatment of obesity-associated diabetes. *Nature reviews Drug discovery,* 15**,** 639-660.

LAMBERT, C. P., ARCHER, R. L. & EVANS, W. J. 2002. Body composition in ambulatory women with multiple sclerosis. *Archives of physical medicine and rehabilitation,* 83**,** 1559-1561.

LAUFS, U., ENDRES, M., STAGLIANO, N., AMIN-HANJANI, S., CHUI, D.-S., YANG, S.-X., SIMONCINI, T., YAMADA, M., RABKIN, E. & ALLEN, P. G. 2000. Neuroprotection mediated by changes in the endothelial actin cytoskeleton. *The Journal of clinical investigation,* 106**,** 15-24.

LAUFS, U. & LIAO, J. K. 1998. Post-transcriptional regulation of endothelial nitric oxide synthase mRNA stability by Rho GTPase. *Journal of Biological Chemistry,* 273**,** 24266-24271.

LEDDY, S. & DOBSON, R. 2020. Multiple sclerosis. *Medicine,* 48**,** 588-594.

LEE, H.-G., CHO, M.-Z. & CHOI, J.-M. 2020a. Bystander CD4+ T cells: crossroads between innate and adaptive immunity. *Experimental & Molecular Medicine,* 52**,** 1255-1263.

LEE, S. H., LEE, J.-Y., HUI KIM, G., JUNG, K. J., LEE, S., KIM, H. C. & JEE, S. H. 2020b. Two-sample Mendelian randomization study of lipid levels and ischemic heart disease. *Korean Circulation Journal,* 50**,** 940-948.

LEI, X., CAI, S., CHEN, Y., CUI, J., WANG, Y., LI, Z. & LI, Y. 2017. Down-regulation of interleukin 7 receptor (IL-7R) contributes to central nervous system demyelination. *Oncotarget,* 8**,** 28395.

LI, H., CHEN, L., MA, X., CUI, P., LANG, W. & HAO, J. 2019. Shared gene expression between multiple sclerosis and ischemic stroke. *Frontiers in Genetics*, 598.
LIAO, J. K. 2002. Isoprenoids as mediators of the biological effects of statins. *The Journal of clinical investigation,* 110**,** 285-288.
LIAO, K., CHEN, K., BRETT, S., GEHMAN, A., SCHWARTZ, A. M., GUNN, G. R. & DEWALL, S. L. 2021. Characterization of the robust humoral immune response to GSK2618960, a humanized anti-IL-7 receptor monoclonal antibody, observed in healthy subjects in a Phase 1 study. *Plos one,* 16**,** e0249049.
LIBBEY, J. E., MCCOY, L. L. & FUJINAMI, R. S. 2007. Molecular mimicry in multiple sclerosis. *International review of neurobiology,* 79**,** 127-147.
LIU, G., SHI, M., MOSLEY, J. D., WENG, C., ZHANG, Y., LEE, M. T. M., JARVIK, G. P., HAKONARSON, H., NAMJOU-KHALES, B. & SLEIMAN, P. 2021a. A Mendelian Randomization Approach Using 3-HMG-Coenzyme-A Reductase Gene Variation to Evaluate the Association of Statin-Induced Low-Density Lipoprotein Cholesterol Lowering With Noncardiovascular Disease Phenotypes. *JAMA network open,* 4**,** e2112820-e2112820.
LIU, Y., LI, B., MA, Y., HUANG, Y., OUYANG, F. & LIU, Q. 2021b. Mendelian Randomization Integrating GWAS, eQTL, and mQTL Data Identified Genes Pleiotropically Associated With Atrial Fibrillation. *Frontiers in Cardiovascular Medicine,* 8.
LU, H., WU, P.-F., ZHANG, W. & XIA, K. 2020. Coffee consumption is not associated with risk of multiple sclerosis: A mendelian randomization study. *Multiple Sclerosis and Related Disorders,* 44**,** 102300.
LUBLIN, F. D., REINGOLD, S. C., COHEN, J. A., CUTTER, G. R., SØRENSEN, P. S., THOMPSON, A. J., WOLINSKY, J. S., BALCER, L. J., BANWELL, B. & BARKHOF, F. 2014. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology,* 83**,** 278-286.
LUNDSTRÖM, W., HIGHFILL, S., WALSH, S. T., BEQ, S., MORSE, E., KOCKUM, I., ALFREDSSON, L., OLSSON, T., HILLERT, J. & MACKALL, C. L. 2013. Soluble IL7Rα potentiates IL-7 bioactivity and promotes autoimmunity. *Proceedings of the National Academy of Sciences,* 110**,** E1761-E1770.
MACARTHUR, J., BOWLER, E., CEREZO, M., GIL, L., HALL, P., HASTINGS, E., JUNKINS, H., MCMAHON, A., MILANO, A. & MORALES, J. 2017. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic acids research,* 45**,** D896-D901.
MAGYARI, M. & SORENSEN, P. S. 2020. Comorbidity in multiple sclerosis. *Frontiers in Neurology,* 11**,** 851.
MALIK, R., CHAUHAN, G., TRAYLOR, M., SARGURUPREMRAJ, M., OKADA, Y., MISHRA, A., RUTTEN-JACOBS, L., GIESE, A.-K., VAN DER LAAN, S. W. & GRETARSDOTTIR, S. 2018. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nature genetics,* 50**,** 524-537.
MANOUCHEHRINIA, A., KINGWELL, E., ZHU, F., TREMLETT, H., HILLERT, J. & RAMANUJAM, R. 2021. A multiple sclerosis disease progression measure based on cumulative disability. *Multiple Sclerosis Journal,* 27**,** 1875-1883.
MANOUCHEHRINIA, A., WESTERLIND, H., KINGWELL, E., ZHU, F., CARRUTHERS, R., RAMANUJAM, R., BAN, M., GLASER, A., SAWCER, S. & TREMLETT, H. 2017. Age related multiple sclerosis severity score: disability ranked by age. *Multiple Sclerosis Journal,* 23**,** 1938-1946.
MANSILLA, M., PRESAS-RODRÍGUEZ, S., TENIENTE-SERRA, A., GONZÁLEZ-LARREATEGUI, I., QUIRANT-SÁNCHEZ, B., FONDELLI, F., DJEDOVIC, N., IWASZKIEWICZ-GRZEŚ, D., CHWOJNICKI, K. & MILJKOVIĆ, Đ. 2021. Paving the way towards an effective treatment for multiple sclerosis: advances in cell therapy. *Cellular & Molecular Immunology,* 18**,** 1353-1374.
MARANS, R. W. 2003. Understanding environmental quality through quality of life studies: the 2001 DAS and its use of subjective and objective indicators. *Landscape and Urban Planning,* 65**,** 73-83.
MARINI, S., MERINO, J., MONTGOMERY, B. E., MALIK, R., SUDLOW, C. L., DICHGANS, M., FLOREZ, J. C., ROSAND, J., GILL, D. & ANDERSON, C. D. 2020. Mendelian randomization study of obesity and cerebrovascular disease. *Annals of neurology,* 87**,** 516-524.
MARRIE, R. A. 2017. Comorbidity in multiple sclerosis: implications for patient care. *Nature Reviews Neurology,* 13**,** 375-382.
MARRIE, R. A., COHEN, J., STUVE, O., TROJANO, M., SØRENSEN, P. S., REINGOLD, S., CUTTER, G. & REIDER, N. 2015. A systematic review of the incidence and prevalence of comorbidity in multiple sclerosis: overview. *Multiple Sclerosis Journal,* 21**,** 263-281.

MARRIE, R. A. & HORWITZ, R. I. 2010. Emerging effects of comorbidities on multiple sclerosis. *The Lancet Neurology,* 9**,** 820-828.
MARTINELLI, V., RODEGHER, M., MOIOLA, L. & COMI, G. 2004. Late onset multiple sclerosis: clinical characteristics, prognostic factors and differential diagnosis. *Neurological Sciences,* 25**,** s350-s355.
MAZZUCCHELLI, R. I., RIVA, A. & DURUM, S. K. The human IL-7 receptor gene: deletions, polymorphisms and mutations. Seminars in immunology, 2012. Elsevier, 225-230.
MCINNES, L., HEALY, J. & MELVILLE, J. 2018. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.
MCTAGGART, F. & JONES, P. 2008. Effects of statins on high-density lipoproteins: a potential contribution to cardiovascular benefit. *Cardiovascular drugs and therapy,* 22**,** 321-338.
MENYHART, O., WELTZ, B. & GYŐRFFY, B. 2021. MultipleTesting. com: a tool for life science researchers for multiple hypothesis testing correction. *PloS one,* 16**,** e0245824.
MITCHELL, A. B., COLE, J. W., MCARDLE, P. F., CHENG, Y.-C., RYAN, K. A., SPARKS, M. J., MITCHELL, B. D. & KITTNER, S. J. 2015. Obesity increases risk of ischemic stroke in young adults. *Stroke,* 46**,** 1690-1692.
MITCHELL, R. E., PATERNOSTER, L. & SMITH, G. D. 2018. Mendelian randomization in case only studies: a promising approach to be applied with caution. *American Journal of Cardiology,* 122**,** 2169-2171.
MOHAMMAD, S., AZIZ, R., AL MAHRI, S., MALIK, S. S., HAJI, E., KHAN, A. H., KHATLANI, T. S. & BOUCHAMA, A. 2021. Obesity and COVID-19: what makes obese host so vulnerable? *Immunity & Ageing,* 18**,** 1-10.
MOKRY, L. E., ROSS, S., TIMPSON, N. J., SAWCER, S., DAVEY SMITH, G. & RICHARDS, J. B. 2016. Obesity and multiple sclerosis: a mendelian randomization study. *PLoS medicine,* 13**,** e1002053.
MUNGER, K. L., CHITNIS, T. & ASCHERIO, A. 2009. Body size and risk of MS in two cohorts of US women. *Neurology,* 73**,** 1543-1550.
NASERI, A., NASIRI, E., SAHRAIAN, M. A., DANESHVAR, S. & TALEBI, M. 2021. Clinical features of late-onset multiple sclerosis: A systematic review and meta-analysis. *Multiple Sclerosis and Related Disorders,* 50**,** 102816.
NASL-KHAMENEH, A. M., MIRSHAFIEY, A., MOGHADASI, A. N., SHIRI-SHAHSAVAR, M. R., ESHRAGHIAN, M. R., SHADANIAN, M., ABDOLAHI, M. & SABOOR-YARAGHI, A. A. 2018. Comparison of cytokine expression in multiple sclerosis patients and healthy volunteers. *Acta Medica Iranica***,** 77-83.
NEFLA, M., DARLING, N. J., VAN GIJSEL BONNELLO, M., COHEN, P. & ARTHUR, J. S. C. 2021. Salt inducible kinases 2 and 3 are required for thymic T cell development. *Scientific reports,* 11**,** 1-18.
NELSON, M. R., TIPNEY, H., PAINTER, J. L., SHEN, J., NICOLETTI, P., SHEN, Y., FLORATOS, A., SHAM, P. C., LI, M. J. & WANG, J. 2015. The support of human genetic evidence for approved drug indications. *Nature genetics,* 47**,** 856-860.
NESS, T. L., EWING, J. L., HOGABOAM, C. M. & KUNKEL, S. L. 2006. CCR4 is a key modulator of innate immune responses. *The Journal of immunology,* 177**,** 7531-7539.
NEUHAUS, O., STÜVE, O., ZARNVIL, S. S. & HARTUNG, H.-P. 2004. Are statins a treatment option for multiple sclerosis? *The Lancet Neurology,* 3**,** 369-371.
NIELSEN, N. M., WESTERGAARD, T., FRISCH, M., ROSTGAARD, K., WOHLFAHRT, J., KOCH-HENRIKSEN, N., MELBYE, M. & HJALGRIM, H. 2006. Type 1 diabetes and multiple sclerosis: A Danish population-based cohort study. *Archives of neurology,* 63**,** 1001-1004.
NIU, P.-P., SONG, B., WANG, X. & XU, Y.-M. 2020. Serum Uric Acid Level and Multiple Sclerosis: A Mendelian Randomization Study. *Frontiers in genetics,* 11**,** 254.
OLSSON, T., BARCELLOS, L. F. & ALFREDSSON, L. 2017. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nature Reviews Neurology,* 13**,** 25-36.
OLSSON, T., BOSTER, A., FERNÁNDEZ, Ó., FREEDMAN, M. S., POZZILLI, C., BACH, D., BERKANI, O., MUELLER, M. S., SIDORENKO, T. & RADUE, E.-W. 2014. Oral ponesimod in relapsing–remitting multiple sclerosis: a randomised phase II trial. *Journal of Neurology, Neurosurgery & Psychiatry,* 85**,** 1198-1208.
ON NEUROSCIENCE, F. & DISORDERS, N. S. 2014. Improving and Accelerating Therapeutic Development for Nervous System Disorders: Workshop Summary.
ØSTERGAARD, S. D., MUKHERJEE, S., SHARP, S. J., PROITSI, P., LOTTA, L. A., DAY, F., PERRY, J. R., BOEHME, K. L., WALTER, S. & KAUWE, J. S. 2015. Associations between potentially modifiable risk factors and Alzheimer disease: a Mendelian randomization study. *PLoS medicine,* 12**,** e1001841.

OUCHI, N., PARKER, J. L., LUGUS, J. J. & WALSH, K. 2011. Adipokines in inflammation and metabolic disease. *Nature reviews immunology,* 11**,** 85-97.

PACHECO, Y., ACOSTA-AMPUDIA, Y., MONSALVE, D. M., CHANG, C., GERSHWIN, M. E. & ANAYA, J.-M. 2019. Bystander activation and autoimmunity. *Journal of autoimmunity,* 103**,** 102301.

PARK, Y.-M., MYERS, M. & VIEIRA-POTTER, V. J. 2014. Adipose tissue inflammation and metabolic dysfunction: role of exercise. *Missouri medicine,* 111**,** 65.

PATERNOSTER, L., TILLING, K. & DAVEY SMITH, G. 2017. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges. *PLoS genetics,* 13**,** e1006944.

PENG, H., WU, X., WEN, Y., LIN, J. & GUAN, W. 2022. Myocardial infarction and stroke risks in multiple sclerosis patients: a two-sample Mendelian randomization study. *Multiple Sclerosis and Related Disorders***,** 103501.

PILUTTI, L. A. & MOTL, R. W. 2019. Body composition and disability in people with multiple sclerosis: A dual-energy x-ray absorptiometry study. *Multiple sclerosis and related disorders,* 29**,** 41-47.

PINEDA-TORRA, I., SIDDIQUE, S., WADDINGTON, K. E., FARRELL, R. & JURY, E. C. 2021. Disrupted Lipid Metabolism in Multiple Sclerosis: A Role for Liver X Receptors? *Frontiers in Endocrinology,* 12**,** 345.

PINHEIRO, M. A. L., KOOIJ, G., MIZEE, M. R., KAMERMANS, A., ENZMANN, G., LYCK, R., SCHWANINGER, M., ENGELHARDT, B. & DE VRIES, H. E. 2016. Immune cell trafficking across the barriers of the central nervous system in multiple sclerosis and stroke. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1862**,** 461-471.

POELLMANN, W., STARCK, M., KUEMPFEL, T. & KOEHLER, J. 2015. Changing course of Multiple Sclerosis (MS) after stroke: Excessive disease activity in a MS patient after cerebral infarction (P1. 122). AAN Enterprises.

POLY, T. N., ISLAM, M. M., WALTHER, B. A., YANG, H.-C., WU, C.-C., LIN, M.-C. & LI, Y.-C. 2020. Association between use of statin and risk of dementia: a meta-analysis of observational studies. *Neuroepidemiology,* 54**,** 214-226.

PONTI, F., SANTORO, A., MERCATELLI, D., GASPERINI, C., CONTE, M., MARTUCCI, M., SANGIORGI, L., FRANCESCHI, C. & BAZZOCCHI, A. 2020. Aging and imaging assessment of body composition: from fat to facts. *Frontiers in endocrinology***,** 861.

POPPENSIEKER, K., OTTE, D.-M., SCHÜRMANN, B., LIMMER, A., DRESING, P., DREWS, E., SCHUMAK, B., KLOTZ, L., RAASCH, J. & MILDNER, A. 2012. CC chemokine receptor 4 is required for experimental autoimmune encephalomyelitis by regulating GM-CSF and IL-23 production in dendritic cells. *Proceedings of the National Academy of Sciences,* 109**,** 3897-3902.

PORCU, E., RÜEGER, S., LEPIK, K., SANTONI, F. A., REYMOND, A. & KUTALIK, Z. 2019. Mendelian randomization integrating GWAS and eQTL data reveals genetic determinants of complex and clinical traits. *Nature communications,* 10**,** 1-12.

PRADHAN, R., NGO, P. A., MARTÍNEZ-SÁNCHEZ, L. D. C., NEURATH, M. F. & LÓPEZ-POSADAS, R. 2021. Rho GTPases as Key Molecular Players within Intestinal Mucosa and GI Diseases. *Cells,* 10**,** 66.

PROITSI, P., LUPTON, M. K., VELAYUDHAN, L., NEWHOUSE, S., FOGH, I., TSOLAKI, M., DANIILIDOU, M., PRITCHARD, M., KLOSZEWSKA, I. & SOININEN, H. 2014. Genetic predisposition to increased blood cholesterol and triglyceride lipid levels and risk of Alzheimer disease: a Mendelian randomization analysis. *PLoS medicine,* 11**,** e1001713.

RÁDIKOVÁ, Ž., PENESOVÁ, A., VLČEK, M., HAVRANOVÁ, A., SIVÁKOVÁ, M., ŠIARNIK, P., ŽITŇANOVÁ, I., IMRICH, R., TURČÁNI, P. & KOLLÁR, B. 2020. Lipoprotein profiling in early multiple sclerosis patients: effect of chronic inflammation? *Lipids in health and disease,* 19**,** 1-10.

RAMASWAMY, M., DUMONT, C., CRUZ, A. C., MUPPIDI, J. R., GOMEZ, T. S., BILLADEAU, D. D., TYBULEWICZ, V. L. & SIEGEL, R. M. 2007. Cutting edge: Rac GTPases sensitize activated T cells to die via Fas. *The Journal of Immunology,* 179**,** 6384-6388.

RAO, V. S. & KAKKAR, V. V. 2011. Friend turns foe: transformation of anti-inflammatory HDL to proinflammatory HDL during acute-phase response. *Cholesterol,* 2011.

REAY, W. R. & CAIRNS, M. J. 2021. Advancing the use of genome-wide association studies for drug repurposing. *Nature Reviews Genetics,* 22**,** 658-671.

REES, J. M., WOOD, A. M., DUDBRIDGE, F. & BURGESS, S. 2019. Robust methods in Mendelian randomization via penalization of heterogeneous causal estimates. *PloS one,* 14**,** e0222362.

REVEZ, J. A., LIN, T., QIAO, Z., XUE, A., HOLTZ, Y., ZHU, Z., ZENG, J., WANG, H., SIDORENKO, J. & KEMPER, K. E. 2020. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. *Nature communications,* 11**,** 1-12.

RHEAD, B., BÄÄRNHIELM, M., GIANFRANCESCO, M., MOK, A., SHAO, X., QUACH, H., SHEN, L., SCHAEFER, C., LINK, J. & GYLLENBERG, A. 2016. Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk. *Neurology Genetics,* 2.

RICHTER, B., CUTTER, G., PANDEY, K., MOHN, J., BACON, T. & KISTER, I. 2017. Body mass index correlates with multiple sclerosis disease and symptom severity in women, but not in men. *Neurol Disord Therap,* 1**,** 1-5.

RIDLEY, A. J. 2006. Rho GTPases and actin dynamics in membrane protrusions and vesicle trafficking. *Trends in cell biology,* 16**,** 522-529.

RIKITAKE, Y. & LIAO, J. K. 2005. Rho GTPases, statins, and nitric oxide. *Circulation research,* 97**,** 1232-1235.

RITCHIE, D. & COLONNA, M. 2018. Mechanisms of action and clinical development of elotuzumab. *Clinical and translational science,* 11**,** 261.

ROXBURGH, R., SEAMAN, S., MASTERMAN, T., HENSIEK, A., SAWCER, S., VUKUSIC, S., ACHITI, I., CONFAVREUX, C., COUSTANS, M. & LE PAGE, E. 2005. Multiple Sclerosis Severity Score: using disability and disease duration to rate disease severity. *Neurology,* 64**,** 1144-1151.

RUDRAPAL, M., KHAIRNAR, S. J. & JADHAV, A. G. 2020. Drug repurposing (DR): an emerging approach in drug discovery. *Drug Repurposing Hypothesis Mol. Asp. Ther. Appl*.

SANDERSON, E. 2021. Multivariable Mendelian randomization and mediation. *Cold Spring Harbor perspectives in medicine,* 11**,** a038984.

SAOUDI, A., KASSEM, S., DEJEAN, A. S. & GAUD, G. 2014. Rho-GTPases as key regulators of T lymphocyte biology. *Small GTPases,* 5**,** e983862.

SAWAYA, A. P., JOZIC, I., STONE, R. C., PASTAR, I., EGGER, A. N., STOJADINOVIC, O., GLINOS, G. D., KIRSNER, R. S. & TOMIC-CANIC, M. 2019. Mevastatin promotes healing by targeting caveolin-1 to restore EGFR signaling. *JCI insight,* 4.

SAWCER, S., HELLENTHAL, G., PIRINEN, M., SPENCER, C. C., PATSOPOULOS, N. A., MOUTSIANAS, L., DILTHEY, A., SU, Z., FREEMAN, C. & HUNT, S. E. 2011. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature,* 476**,** 214.

SCHAEFFER, J., COSSETTI, C., MALLUCCI, G. & PLUCHINO, S. 2015a. Chapter 30 - Multiple Sclerosis. *In:* ZIGMOND, M. J., ROWLAND, L. P. & COYLE, J. T. (eds.) *Neurobiology of Brain Disorders.* San Diego: Academic Press.

SCHAEFFER, J., COSSETTI, C., MALLUCCI, G. & PLUCHINO, S. Chapter 30@ Multiple Sclerosis. 2015b.

SCHEU, S., ALI, S., RULAND, C., AROLT, V. & ALFERINK, J. 2017. The CC chemokines CCL17 and CCL22 and their receptor CCR4 in CNS autoimmunity. *International journal of molecular sciences,* 18**,** 2306.

SCHORR, M., DICHTEL, L. E., GERWECK, A. V., VALERA, R. D., TORRIANI, M., MILLER, K. K. & BREDELLA, M. A. 2018. Sex differences in body composition and association with cardiometabolic risk. *Biology of sex differences,* 9**,** 1-10.

SELVIN, E., PAYNTER, N. P. & ERLINGER, T. P. 2007. The effect of weight loss on C-reactive protein: a systematic review. *Archives of internal medicine,* 167**,** 31-39.

SENA, A., PEDROSA, R. & MORAIS, M. G. 2003. Therapeutic potential of lovastatin in multiple sclerosis. *Journal of neurology***,** 754-755.

SHEEHAN, N. A. & DIDELEZ, V. 2011. Commentary: Can 'many weak'instruments ever be 'strong'? *International journal of epidemiology,* 40**,** 752-754.

SHI, K., TIAN, D.-C., LI, Z.-G., DUCRUET, A. F., LAWTON, M. T. & SHI, F.-D. 2019. Global brain inflammation in stroke. *The Lancet Neurology,* 18**,** 1058-1066.

SHIMIZU, Y., OTA, K., KUBO, S., KABASAWA, C., KOBAYASHI, M., OHASHI, T. & UCHIYAMA, S. 2011. Association of Th1/Th2-related chemokine receptors in peripheral T cells with disease activity in patients with multiple sclerosis and neuromyelitis optica. *European neurology,* 66**,** 91-97.

SIMMONS, D. P., NGUYEN, H. N., GOMEZ-RIVAS, E., JEONG, Y., JONSSON, A. H., CHEN, A. F., LANGE, J. K., DYER, G. S., BLAZAR, P., EARP, B. E., COBLYN, J. S., MASSAROTTI, E. M., SPARKS, J. A., TODD, D. J., RAO, D. A., KIM, E. Y. & BRENNER, M. B. 2022. SLAMF7 engagement superactivates macrophages in acute and chronic inflammation. *Science Immunology,* 7**,** eabf2846.

SIMPSON, S., BLIZZARD, L., OTAHAL, P., VAN DER MEI, I. & TAYLOR, B. 2011. Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. *Journal of Neurology, Neurosurgery & Psychiatry,* 82**,** 1132-1141.

SIRONI, M., GUERINI, F. R., AGLIARDI, C., BIASIN, M., CAGLIANI, R., FUMAGALLI, M., CAPUTO, D., CASSINOTTI, A., ARDIZZONE, S. & ZANZOTTERA, M. 2011. An evolutionary analysis of RAC2 identifies haplotypes associated with human autoimmune diseases. *Molecular biology and evolution,* 28**,** 3319-3329.

SMITH, G. D., TIMPSON, N. & EBRAHIM, S. 2008. Strengthening causal inference in cardiovascular epidemiology through Mendelian randomization. *Annals of medicine,* 40**,** 524-541.

SOELBERG SORENSEN, P. 2017. Safety concerns and risk management of multiple sclerosis therapies. *Acta Neurologica Scandinavica,* 136**,** 168-186.

SOLDÁN, M. M. P., PITTOCK, S. J., WEIGAND, S. D., YAWN, B. P. & RODRIGUEZ, M. 2012. Statin therapy and multiple sclerosis disability in a population-based cohort. *Multiple Sclerosis Journal,* 18**,** 358-363.

SONEHARA, K. & OKADA, Y. 2021. Genomics-driven drug discovery based on disease-susceptibility genes. *Inflammation and regeneration,* 41**,** 1-5.

SORENSEN, P. S., LISBY, S., GROVE, R., DEROSIER, F., SHACKELFORD, S., HAVRDOVA, E., DRULOVIC, J. & FILIPPI, M. 2014. Safety and efficacy of ofatumumab in relapsing-remitting multiple sclerosis: a phase 2 study. *Neurology,* 82**,** 573-581.

SPARKS, D. L., KRYSCIO, R. J., SABBAGH, M. N., CONNOR, D. J., SPARKS, L. M. & LIEBSACK, C. 2008. Reduced risk of incident AD with elective statin use in a clinical trial cohort. *Current Alzheimer Research,* 5**,** 416-421.

SPEED, M. S., JEFSEN, O. H., BØRGLUM, A. D., SPEED, D. & ØSTERGAARD, S. D. 2019. Investigating the association between body fat and depression via Mendelian randomization. *Translational psychiatry,* 9**,** 1-9.

STORM, C. S., KIA, D. A., ALMRAMHI, M. M., BANDRES-CIGA, S., FINAN, C., HINGORANI, A. D. & WOOD, N. W. 2021. Finding genetically-supported drug targets for Parkinson’s disease using Mendelian randomization of the druggable genome. *Nature communications,* 12**,** 1-14.

SUDLOW, C., GALLACHER, J., ALLEN, N., BERAL, V., BURTON, P., DANESH, J., DOWNEY, P., ELLIOTT, P., GREEN, J. & LANDRAY, M. 2015. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine,* 12**,** e1001779.

SULTANA, J., CRISAFULLI, S., GABBAY, F., LYNN, E., SHAKIR, S. & TRIFIRÒ, G. 2020. Challenges for drug repurposing in the COVID-19 pandemic era. *Frontiers in pharmacology***,** 1657.

SUN, B. B., MARANVILLE, J. C., PETERS, J. E., STACEY, D., STALEY, J. R., BLACKSHAW, J., BURGESS, S., JIANG, T., PAIGE, E. & SURENDRAN, P. 2018. Genomic atlas of the human plasma proteome. *Nature,* 558**,** 73-79.

SUN, Z., JIANG, Q., LI, J. & GUO, J. 2020. The potent roles of salt-inducible kinases (SIKs) in metabolic homeostasis and tumorigenesis. *Signal transduction and targeted therapy,* 5**,** 1-15.

TADIĆ, D., ĐAJIĆ, V., GRGIĆ, S. & MILJKOVIĆ, S. 2020. Association of body mass index with progression and prediction of multiple sclerosis. *Scripta Medica,* 51**,** 34-40.

TAKEMOTO, M. & LIAO, J. K. 2001. Pleiotropic effects of 3-hydroxy-3-methylglutaryl coenzyme a reductase inhibitors. *Arteriosclerosis, thrombosis, and vascular biology,* 21**,** 1712-1719.

TAO, C., SIMPSON, S., VAN DER MEI, I., BLIZZARD, L., HAVRDOVA, E., HORAKOVA, D., SHAYGANNEJAD, V., LUGARESI, A., IZQUIERDO, G. & TROJANO, M. 2016. Higher latitude is significantly associated with an earlier age of disease onset in multiple sclerosis. *J Neurol Neurosurg Psychiatry***,** jnnp-2016-314013.

TEAM, R. C. 2017. R Foundation for Statistical Computing; Vienna, Austria: 2016. R: A language and environment for statistical computing. *URL http://www. R-project. org/[Google Scholar]*.

TELL, R. M., KIMURA, K. & PALIĆ, D. 2012. Rac2 expression and its role in neutrophil functions of zebrafish (Danio rerio). *Fish & shellfish immunology,* 33**,** 1086-1094.

TETTEY, P., SIMPSON JR, S., TAYLOR, B., BLIZZARD, L., PONSONBY, A.-L., DWYER, T., KOSTNER, K. & VAN DER MEI, I. 2014. An adverse lipid profile is associated with disability and progression in disability, in people with MS. *Multiple Sclerosis Journal,* 20**,** 1737-1744.

THOMPSON A, B. S., GEURTS J, HEMMER B, CICCARELLI O 2018. Multiple sclerosis. *Lancet,* 391**,** 1622–1636.

THORMANN, A., MAGYARI, M., KOCH-HENRIKSEN, N., LAURSEN, B. & SØRENSEN, P. S. 2016. Vascular comorbidities in multiple sclerosis: a nationwide study from Denmark. *Journal of neurology,* 263**,** 2484-2493.

TIAN, Z., SONG, Y., YAO, Y., GUO, J., GONG, Z. & WANG, Z. 2020a. Genetic etiology shared by multiple sclerosis and ischemic stroke. *Frontiers in genetics,* 11**,** 646.
TIAN, Z., SONG, Y., YAO, Y., GUO, J., GONG, Z. & WANG, Z. 2020b. Genetic etiology shared by multiple sclerosis and ischemic stroke. *Frontiers in genetics***,** 646.
TRAPP, B. D., PETERSON, J., RANSOHOFF, R. M., RUDICK, R., MÖRK, S. & BÖ, L. 1998. Axonal transection in the lesions of multiple sclerosis. *New England Journal of Medicine,* 338**,** 278-285.
TSENG, C. H., HUANG, W. S., LIN, C. L. & CHANG, Y. J. 2015. Increased risk of ischaemic stroke among patients with multiple sclerosis. *European Journal of Neurology,* 22**,** 500-506.
TWORK, S., WIESMETH, S., SPINDLER, M., WIRTZ, M., SCHIPPER, S., PÖHLAU, D., KLEWER, J. & KUGLER, J. 2010. Disability status and quality of life in multiple sclerosis: non-linearity of the Expanded Disability Status Scale (EDSS). *Health and quality of life outcomes,* 8**,** 1-6.
VANDENBROECK, K. 2012. Cytokine gene polymorphisms and human autoimmune disease in the era of genome-wide association studies. *Journal of Interferon & Cytokine Research,* 32**,** 139-151.
VANDERWEELE, T. J., TCHETGEN, E. J. T., CORNELIS, M. & KRAFT, P. 2014. Methodological challenges in mendelian randomization. *Epidemiology (Cambridge, Mass.),* 25**,** 427.
VIJAYAN, M., KUMAR, S., BHATTI, J. & REDDY, P. 2017. Molecular links and biomarkers of stroke, vascular dementia, and Alzheimer's disease. *Progress in molecular biology and translational science,* 146**,** 95-126.
VÕSA, U., CLARINGBOULD, A., WESTRA, H.-J., BONDER, M. J., DEELEN, P., ZENG, B., KIRSTEN, H., SAHA, A., KREUZHUBER, R. & BRUGGE, H. 2021. Large-scale cis-and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nature genetics,* 53**,** 1300-1310.
WALKER, V. M., DAVEY SMITH, G., DAVIES, N. M. & MARTIN, R. M. 2017. Mendelian randomization: a novel approach for the prediction of adverse drug events and drug repurposing opportunities. *International journal of epidemiology,* 46**,** 2078-2089.
WALTON, C., KING, R., RECHTMAN, L., KAYE, W., LERAY, E., MARRIE, R. A., ROBERTSON, N., LA ROCCA, N., UITDEHAAG, B. & VAN DER MEI, I. 2020. Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS. *Multiple Sclerosis Journal,* 26**,** 1816-1821.
WANG, C.-Y., LIU, P.-Y. & LIAO, J. K. 2008. Pleiotropic effects of statin therapy: molecular mechanisms and clinical results. *Trends in molecular medicine,* 14**,** 37-44.
WANG, D., LIU, S., WARRELL, J., WON, H., SHI, X., NAVARRO, F. C., CLARKE, D., GU, M., EMANI, P. & YANG, Y. T. 2018. Comprehensive functional genomic resource and integrative model for the human brain. *Science,* 362**,** eaat8464.
WANG, J., ZHAO, Q., BOWDEN, J., HEMANI, G., SMITH, G. D., SMALL, D. S. & ZHANG, N. R. 2021. Causal inference for heritable phenotypic risk factors using heterogeneous genetic instruments. *PLoS genetics,* 17**,** e1009575.
WANG, W., OLSSON, T., KOSTULAS, V., HÖJEBERG, B., EKRE, H. & LINK, H. 1992. Myelin antigen reactive T cells in cerebrovascular diseases. *Clinical & Experimental Immunology,* 88**,** 157-162.
WATSON, S. & MARX, J. B. 2019. Mogamulizumab-kpkc: A novel therapy for the treatment of cutaneous T-cell lymphoma. *Journal of the Advanced Practitioner in Oncology,* 10**,** 883.
WEBER, M. S., STEINMAN, L. & ZAMVIL, S. S. 2007. Statins—treatment option for central nervous system autoimmune disease? *Neurotherapeutics,* 4**,** 693-700.
WEI, T., SIMKO, V., LEVY, M., XIE, Y., JIN, Y. & ZEMLA, J. 2017. Package 'corrplot'. *Statistician,* 56**,** e24.
WEINSTOCK-GUTTMAN, B., ZIVADINOV, R., MAHFOOZ, N., CARL, E., DRAKE, A., SCHNEIDER, J., TETER, B., HUSSEIN, S., MEHTA, B. & WEISKOPF, M. 2011. Serum lipid profiles are associated with disability and MRI outcomes in multiple sclerosis. *Journal of neuroinflammation,* 8**,** 127.
WEISSGLAS-VOLKOV, D. & PAJUKANTA, P. 2010. Genetic causes of high and low serum HDL-cholesterol. *Journal of lipid research,* 51**,** 2032-2057.
WILLER, C. J., SCHMIDT, E. M., SENGUPTA, S., PELOSO, G. M., GUSTAFSSON, S., KANONI, S., GANNA, A., CHEN, J., BUCHKOVICH, M. L. & MORA, S. 2013. Discovery and refinement of loci associated with lipid levels. *Nature genetics,* 45**,** 1274.
WILLIAMS, D. M., FINAN, C., SCHMIDT, A. F., BURGESS, S. & HINGORANI, A. D. 2020. Lipid lowering and Alzheimer disease risk: a mendelian randomization study. *Annals of neurology,* 87**,** 30-39.

WILT, T. J., BLOOMFIELD, H. E., MACDONALD, R., NELSON, D., RUTKS, I., HO, M., LARSEN, G., MCCALL, A., PINEROS, S. & SALES, A. 2004. Effectiveness of statin therapy in adults with coronary heart disease. *Archives of internal medicine,* 164**,** 1427-1436.

WINGO, B. C., YOUNG, H.-J. & MOTL, R. W. 2018. Body composition differences between adults with multiple sclerosis and BMI-matched controls without MS. *Disability and health journal,* 11**,** 243-248.

WINSLØW, U. C., RODE, L. & NORDESTGAARD, B. G. 2015. High tobacco consumption lowers body weight: a Mendelian randomization study of the Copenhagen General Population Study. *International journal of epidemiology,* 44**,** 540-550.

WOLINSKY, J. S., NARAYANA, P. A., O'CONNOR, P., COYLE, P. K., FORD, C., JOHNSON, K., MILLER, A., PARDO, L., KADOSH, S. & LADKANI, D. 2007. Glatiramer acetate in primary progressive multiple sclerosis: results of a multinational, multicenter, double-blind, placebo-controlled trial. *Annals of neurology,* 61**,** 14-24.

WOLOZIN, B., KELLMAN, W., RUOSSEAU, P., CELESIA, G. G. & SIEGEL, G. 2000. Decreased prevalence of Alzheimer disease associated with 3-hydroxy-3-methyglutaryl coenzyme A reductase inhibitors. *Archives of neurology,* 57**,** 1439-1443.

WOOD, H. 2017. Multiple sclerosis: latitude and vitamin D influence disease course in multiple sclerosis. *Nature reviews. Neurology,* 13**,** 3.

YANG, G. & SCHOOLING, C. M. 2021. Investigating genetically mimicked effects of statins via HMGCR inhibition on immune-related diseases in men and women using Mendelian randomization. *Scientific reports,* 11**,** 1-7.

YANG, Q., PAN, W. & QIAN, L. 2017. Identification of the miRNA–mRNA regulatory network in multiple sclerosis. *Neurological research,* 39**,** 142-151.

YAO, C., CHEN, G., SONG, C., KEEFE, J., MENDELSON, M., HUAN, T., SUN, B. B., LASER, A., MARANVILLE, J. C. & WU, H. 2018. Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease. *Nature communications,* 9**,** 1-11.

YOUSUFUDDIN, M. & YOUNG, N. 2019. Aging and ischemic stroke. *Aging (Albany NY),* 11**,** 2542.

YUAN, S., XIONG, Y. & LARSSON, S. C. 2021. An atlas on risk factors for multiple sclerosis: a Mendelian randomization study. *Journal of Neurology,* 268**,** 114-124.

YUKSEL, B., KOC, P., GOKSU, E. O., KARACAY, E., KURTULUS, F., CEKIN, Y. & GOMCELI, Y. B. 2019. Is multiple sclerosis a risk factor for atherosclerosis? *Journal of Neuroradiology*.

ZAHID, M., BUSMAIL, A., PENUMETCHA, S. S., AHLUWALIA, S., IRFAN, R., KHAN, S. A., REDDY, S. R., LOPEZ, M. E. V. & MOHAMMED, L. 2021. Tumor Necrosis Factor Alpha Blockade and Multiple Sclerosis: Exploring New Avenues. *Cureus,* 13.

ZHANG, T., TREMLETT, H., ZHU, F., KINGWELL, E., FISK, J. D., BHAN, V., CAMPBELL, T., STADNYK, K., CARRUTHERS, R. & WOLFSON, C. 2018. Effects of physical comorbidities on disability progression in multiple sclerosis. *Neurology,* 90**,** e419-e427.

ZHENG, J., BAIRD, D., BORGES, M.-C., BOWDEN, J., HEMANI, G., HAYCOCK, P., EVANS, D. M. & SMITH, G. D. 2017. Recent Developments in Mendelian Randomization Studies. *Current epidemiology reports,* 4**,** 330-345.

ZHENG, J., BRION, M. J., KEMP, J. P., WARRINGTON, N. M., BORGES, M. C., HEMANI, G., RICHARDSON, T. G., RASHEED, H., QIAO, Z. & HAYCOCK, P. 2020. The Effect of Plasma Lipids and Lipid-Lowering Interventions on Bone Mineral Density: A Mendelian Randomization Study. *Journal of Bone and Mineral Research,* 35**,** 1224-1235.

ZHENG, J., FRYSZ, M., KEMP, J. P., EVANS, D. M., DAVEY SMITH, G. & TOBIAS, J. H. 2019. Use of Mendelian randomization to examine causal inference in osteoporosis. *Frontiers in endocrinology,* 10**,** 807.

ZHORNITSKY, S., MCKAY, K. A., METZ, L. M., TEUNISSEN, C. E. & RANGACHARI, M. 2016. Cholesterol and markers of cholesterol turnover in multiple sclerosis: relationship with disease outcomes. *Multiple sclerosis and related disorders,* 5**,** 53-65.

ZHOU, Y., CUI, C., MA, X., LUO, W., ZHENG, S. G. & QIU, W. 2020. Nuclear factor κB (NF-κB)–mediated inflammation in multiple sclerosis. *Frontiers in immunology,* 11**,** 391.

ZHU, Y., CASEY, P. J., KUMAR, A. P. & PERVAIZ, S. 2013. Deciphering the signaling networks underlying simvastatin-induced apoptosis in human cancer cells: evidence for non-canonical activation of RhoA and Rac1 GTPases. *Cell death & disease,* 4**,** e568-e568.

ZÖLLER, B., LI, X., SUNDQUIST, J. & SUNDQUIST, K. 2012. Risk of subsequent ischemic and hemorrhagic stroke in patients hospitalized for immune-mediated diseases: a nationwide follow-up study from Sweden. *BMC neurology,* 12**,** 1-9.