Using Mendelian randomization to elucidate the role of vitamin D in multiple sclerosis

Lauren E Mokry

Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal

April 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science.

Table of Contents:

Preface

Author Contributions

Lauren Mokry completed the literature reviews, the statistical analyses and drafting of the manuscript published in PLOS Medicine as well as this thesis. Dr. Richards offered supervision and guidance of immeasurable importance for these projects. Dr. Richards also conceived the design of the experiment published in PLOS Medicine and assisted in drafting of the manuscript. Drs. Thanassoulis, Greenwood, Forgetta and Ross provided important intellectual guidance and revisions to the paper. Drs. Ahmad, Smith and Leong also helped in revising the PLOS Medicine paper. Lastly, Drs. Richards and Ross helped to revise this thesis. Maxime Turgeon graciously translated the abstract of this thesis into French.

Acknowledgement

I would like to thank my supervisor Dr. Brent Richards for his endless support and guidance over the past few years. My thesis committee, Drs. Celia Greenwood and George Thanassoulis, who both provided critical feedback. To Drs. Stephanie Ross and Vince Forgetta, and all members of our lab who helped to foster an easy and collaborative working environment, I express my gratitude. Lastly to my parents who have taken a shine to epidemiology since I started my Master's.

I was also like to thank the International Multiple Sclerosis Consortium, in particular Dr. Stephen Sawcer, for access to their data.

My Master’s has been kindly supported by Dr. Richards and the Canadian Institute for Health Research.

Abstract

Background: Observational studies have long reported an association between decreased vitamin D and increased risk of multiple sclerosis (MS). However since it is difficult to fully protect these approaches from confounding and reverse causation, it remains unclear whether vitamin D is a causal risk factor in MS etiology.

Objectives: To review the current literature of the association between vitamin D and MS and discuss the limitations of these previous approaches. Next, to introduce Mendelian randomization (MR) and apply its principles to investigate whether the genetic determinants of vitamin D are associated with MS susceptibility.

Methods: PubMed was used to search for relevant ecological, observational and randomized controlled trials. For our MR analysis, we first selected single nucleotide polymorphisms (SNPs) that achieved genome-wide significant (*p*-value < 5 x10^{-8}) for 25-hydroxyvitamin D (25OHD), the clinical determinant of vitamin D status, in the SUNLIGHT consortium (N= 33,996). We then obtained effect sizes for these SNPs upon MS in the International Multiple Sclerosis Genetics Consortium (IMSGC), the largest genome-wide association study for MS (including up to 14,498 cases and 24,091 controls). MR estimates were obtained by weighting each SNP's effect on MS by its effect on 25OHD, with estimates pooled to provide a summary measure of the effect of genetically lowered vitamin D upon risk of MS.

Results: Results of our MR analysis using four vitamin D associated SNPs, demonstrated that a 1 standard deviation decrease in natural log 25OHD increased odds of MS by 2 fold (OR=2.02, 95% CI: 1.7–2.5; p = 7.7 × 10^{-12}; I^2 = 63%, 95% CI: 0%–88%).

Conclusions: Using MR, our findings support vitamin D as a causal risk factor for MS, substantiating conclusions first suggested by observational analyses. This provides rationale to promote vitamin D awareness among individual at risk for MS. Whether vitamin D supplementation can prevent MS warrants further investigation by long-term clinical trials.

Abrégé

Contexte: Des études observationelles rapportent depuis longtemps une association entre la baisse du niveau de vitamine D et l'augmentation du risque de sclérose en plaques (SP). Or, puisqu'il est difficile de complètement protéger ces approches de l'effet des facteurs de confusion et de la rétrocausalité, il reste toujours à déterminer si le niveau de vitamine D est réellement un facteur de risque causal dans l'étiologie de la SP.

Objectifs: Passer en revue les publications portant sur l'association entre la vitamine D et la SP et discuter des limites des approches précédentes. Ensuite, introduire la randomisation mendélienne (RM) et appliquer ses principes pour évaluer si les déterminants génétiques de la vitamine D sont associés à la prédisposition à la SP.

Méthodes: PubMed a été utilisé pour identifier les études randomisées, observationelles et écologiques pertinentes. Pour notre analyse de RM, nous avons d'abord sélectionné les polymorphismes nucléotidiques (SNP, single-nucleotide polymorphism) qui étaient significatifs au niveau du génome (valeur-p $< 5 \times 10^{-8}$) pour la 25-hydroxyvitamine D (25OHD), le déterminant clinique du niveau de vitamine D, dans l'étude du consortium SUNLIGHT (N = 33,996). Nous avons ensuite obtenu les tailles d'effet des SNPs sur la SP pour l'étude du International Multiple Sclerosis Genetics Consortium (IMSGC), la plus importante étude d'association pangénomique portant sur la SP (incluant 14,498 cas et 24,091 témoins). Les estimés de l'étude de RM ont ensuite été obtenus en pondérant l'effet de chaque SNP sur la SP par son effet sur la 25OHD, les estimés étant finalement groupés afin d'obtenir une mesure conjointe de l'effet sur le risque de SP d'un niveau génétiquement bas de vitamine D.

Résultats: Les résultats de notre analyse de RM utilisant quatre SNPs associés à la vitamine D démontrent qu'une baisse du niveau de 25OHD équivalant à une erreur type (sur l'échelle logarithmique) double la cote de la SP (OR=2.02, 95% CI: 1.7–2.5; p = 7.7 $\times 10^{-12}$; I2 = 63%, 95% CI: 0%–88%).

Conclusions: En utilisant la RM, nos conclusions supportent l'hypothèse de la vitamine D comme facteur de risque causal pour la SP, appuyant les conclusions d'abord suggérées par les études observationelles. Elles supportent aussi la promotion de la vitamine D chez les individus à risque de SP. Toutefois, des études randomisées de longue haleine sont nécessaires pour déterminer si les suppléments de vitamine D peuvent prévenir la SP.

Chapter 1 – Introduction

Observational studies have reported that decreased levels of vitamin D significantly increase the risk of multiple sclerosis (MS).(1) However, results from randomized controlled trials (RCTs) have been inconsistent.(2) Since the possibility of confounding and reverse causation inherent to observational study designs precludes causal interpretation, these initial associations have not changed clinical guidelines for MS patients.(3) Yet, addressing the role of vitamin D in MS etiology is of great importance from a public health perspective for two main reasons. First, population vitamin D levels have decreased over a 20-year period according to the National Health and Nutrition Examination Survey (NHANES), potentially placing a high proportion of the population at risk for the disease.(4,5) Second, treating MS is costly, with an average course of treatment costing upwards $60,000 annually,(6) whereas vitamin D supplements are both relatively safe and inexpensive, if given at a reasonable dose. Therefore whether vitamin D represents an effective form of prevention for this debilitating condition warrants further investigation. While this question is best addressed using large long-term RCTs, vitamin D is off-patent and therefore the likelihood of an industry-funded trial is small. In the absence of RCT data, Mendelian randomization (MR) can provide insights into the role of biomarkers in disease etiology.

MR has emerged as a powerful study design that can provide evidence supporting, or refuting, causality by utilizing the genetic determinants of a risk factor.(7) Due to the random assortment and segregation of alleles at meiosis, many parallels can be drawn between MR studies and RCTs, where instead of randomization to an intervention group, an individual is randomized to carry a genetic variant.(7) This process of

randomization, distributes all potential confounders equally among allele carriers, ensuring that the genetic variants, or single nucleotide polymorphisms (SNPs), are not associated with these variables and largely free from their bias. Therefore SNPs associated with a risk factor through genome-wide association studies (GWAS) can be used as instruments to investigate the relationship between the risk factor and disease in an un-confounded manner.(7) This is an important advantage since observational analyses involving vitamin D are likely to be confounded, as the self-selected determinants of vitamin D status, such as exercise or a healthy diet, may independently influence MS risk. In addition, since genetic variants are assigned at conception and largely stable over the course of an individual's life, MR studies have the correct temporal ordering to overcome the possibility of reverse causation, with estimates representing *lifetime* risk due to increased or decreased levels of a risk factor. Thus in instances where evidence from trials is lacking, MR studies may represent the best available evidence.

The implementation of MR design is much less expensive than the initiation of clinical trials since it relies upon data previously generated by large GWAS. By employing the two sample MR approach,(8) effect estimates of the selected instruments upon the risk factor and outcome can be obtained from the summary-level statistics of their respective GWAS. This approach can improve power over the traditional MR method by using the full sample size of GWAS instead of relying upon smaller cohorts where both the exposure and outcome were measured.(8) Considering that the sample sizes of most GWAS is well over 10,000 individuals, MR studies are often adequately powered to detect a causal effect. In this approach, MR estimates can be calculated by weighting

each SNP's effect upon the outcome by its effect upon the risk factor, with results pooled using meta-analytic models to provide a summary measure.(8) The implementation of this design has been facilitated in recent years with summary statistics of these large genetic studies being made increasingly available to the public.

MR methods have been used previously to dismiss a causal role for once-prominent risk factor paradigms in cardiovascular disease such as HDL-cholesterol(9) and CRP protein,(10) while confirming others(9,11,12) (see our recent review for a thorough discussion of this topic).(13) Since these MR studies were largely concordant with RCT evidence,(13) MR has demonstrated an ability to identify effective clinical interventions. While the MR design has been predominantly applied to cardiovascular outcomes, in part due to the availability modifiable risk factors and large publicly available datasets, GWAS studies have interrogated the genetic determinants of many outcomes including both vitamin D and MS, providing the possibility to extend this method outside CVD.

Thus the principles of MR can be used to clarify the role of vitamin D in MS etiology – the central aim of this thesis. This thesis is structured so as to first introduce the research question, the current evidence for the association between vitamin D and MS, the limitations of these approaches, the need for additional research and the feasibility of applying MR to this paradigm. Results of our MR analysis are presented within the manuscript provided in Chapter 3 as published in *PLOS Medicine*. Finally, an overall conclusion regarding both our findings and more broadly, the utility of MR to identify effective clinical interventions, is the discussed in Chapter 4.

Chapter 2 – Literature review: Vitamin D and MS Susceptibility

The causes of MS remain poorly characterized, and consequently the disease represents an important unmet clinical need. Since treatment regimens are both costly and somewhat ineffective,(14) it is important to identify causal modifiable risk factors to help guide prevention and drug development efforts. Many studies have reported a significant association between vitamin D and MS.(1,15–17) However the quality of these studies vary, with different measurements for vitamin D and clinical endpoints for MS. This chapter will review the current evidence for this association grouped by study design and discuss the limitations of these previous approaches. Studies included within this section were identified through PubMed and include a measurement of vitamin D (or reliable proxy) as exposure and either a measure of MS occurrence or disease progression as outcome.

Evidence from Ecological Studies

An unequal geographic distribution of MS prevalence and incidence has long been described by ecological studies, with Northern latitude populations exhibiting higher rates of the disease.(18–20) This latitudinal gradient has been replicated in diverse populations,(21,22) and confirmed by a meta-analysis published by Simpson et al. in 2011.(20)

These studies have also suggested that this latitudinal gradient may be mediated by decreased exposure to sunlight,(18,23) and further by vitamin D which is partially derived from sunlight (see **Fig 3.** in Chapter 3).(24) This has been substantiated by the attenuation of this gradient due to proximity to coastal region(25) or vitamin D repletion activities among inhabitants of the Arctic Circle.(26) However, it is difficult to ascertain

the effects of latitude on the risk of MS because this study design is highly susceptible to confounding due to its use of aggregate-level data. For example, this latitudinal gradient may be mediated by ancestry and genetic predisposition or due to the effects of different age-structures across countries. Additionally, since vitamin D levels were not measured at the individual-level, we must make assumptions regarding latitude as a proxy for vitamin D dosage, which may not always prove to be valid. For instance, Northern inhabitants may in reality spend winters in the south, and this individual variability is generally not accounted in the ecological design.(27) Thus while ecological studies can be useful for hypothesis generating, overall the evidence they provide with regards to disease etiology is of poor quality.

Evidence from Observational Studies

Observational studies offer numerous advantages over the ecological design. Here, the effect of the exposure is measured at an individual-level, providing a more accurate estimate of dosage. Additionally, exposure and outcome assessment at the individual-level offers better insight into disease incidence. MS is a rare disease (ranging from 50-200 cases per 100,000 among European populations),(28,29) which poses difficulties for the prospective studies since a large sample size is required to capture a sufficient number of cases. As such, large cohort studies investigating vitamin D and MS incidence have been limited to analyses of the Nurse's Health Study (NHS). Munger et al. first demonstrated a link between vitamin D intake and MS in 187,563 women from the NHS. The study reported a 40% reduced risk of MS among women supplementing with ≥ 400 IU/day of vitamin D relative to women not taking supplements over 10-20 years of follow-up (RR=0.60, 95% CI=0.39-0.92, p-value=0.009) (**Table 1**).(1) A similar

study by the same authors investigating adolescent vitamin D intake in the NHS, produced a suggestive, yet not statistically significant result for this protective effect (RR=0.73, 95% CI=0.50-1.07, p-value=0.18) (N=119,786) (**Table 1**).(30) Notably both studies relied upon self-reported vitamin D intake rather than measuring 25-hydroxyvitamin D (25OHD), the clinical determinant of vitamin D status.

Two small case-control studies using 25OHD measurements provided additional evidence for the association between vitamin D and MS incidence (**Table 1**).(17,31) In particular, a well-designed nested case-control study using data from the Department of Defense Serum Repository (N=148 cases, 296 controls) reported 41% decrease in odds of MS among white individuals for every 50 nmol/L increase in 25OHD (OR=0.59, 95% CI=0.36-0.97, p-value=0.04) (**Table 1**).(17) The case-control design is advantageous when studying a rare disease such as MS, with the nested case control design in particular retaining many attributes of prospective studies. However selection bias can be induced when matching controls to case characteristics, which in turn can affect estimates and limiting generalizability of results.(32) Nonetheless these studies provide critical evidence using 25OHD measurements – a more reliable, yet expensive measure of vitamin D status. In combination with results from the NHS, these analyses provide insight into a possible link between vitamin D and MS incidence.

Others studies have investigated the relationship between vitamin D and disease progression. Compelling evidence came from a study by Ascherio et al. that used multiple 25OHD measurements (at baseline, 6, 12 and 24 months) to assess MS disease activity among participants with clinically isolated syndrome.(15) The study found that a 50 nmol/L increase in 25OHD offered protection from new active lesions

(RR= 0.61, 95% CI=0.44-0.83, p-value=0.002), but did not significantly improve relapse rate or EDSS score (**Table 1**).(33) A study by Mowry et al. involving 469 participants diagnosed with either clinically isolated syndrome or relapse-remitting MS, not only found 25OHD to be protective of new lesions but also improved disability measures (**Table 1**).(34) Two studies found 25OHD status to predict relapse risk among patients with relapse-remitting MS (**Table 1**).(35,36) These studies provide further evidence of vitamin D's involvement in MS etiology.

Other observational studies investigating important predictors of vitamin D status such as skin colour,(37,38) sun exposure,(26,37) month of birth,(39) and oily fish consumption,(26,40) have also provided evidence for this association. In these studies traits expected to decrease vitamin D associated with an increased risk of MS, further recapitulating the direction of effect established in cohort and case-control studies.

While the results from observational studies surely bolster the evidence linking vitamin D to MS, like ecological studies, they possess important limitations that impede any causal interpretation. First, the observational design is susceptible to confounding since exposure to vitamin D is self-selected. Vitamin D is likely to be confounded since individuals who take supplements or ensure sufficiency through diet, are likely to engage in other healthy behaviors that may influence risk of MS. This is known as the healthy user bias.(41) Although many of these studies have adjusted for common confounders, such as age and sex, the potential effect of confounding by unknown or unmeasured variables cannot be ascertained. Additionally, body mass index (BMI), which has been shown to influence both vitamin D(42) and MS,(43,44) was not adjusted

for in most analyses. Therefore, this possibility of confounding could in turn be driving the observed associations.

Small sample size is another weakness of the previous studies. MS is a rare disease, and consequently most studies only analyzed a few hundred MS cases, creating an issue of inadequate power especially when categorizing vitamin D status, as was done in the NHS. Lastly, most studies relied upon self-reported vitamin D intake or other repletion behaviours, which are not as reliable as 25OHD measurements. The ideal, yet costly, study design to investigate this paradigm would be large prospective study with 25OHD measured at multiple intervals.

Evidence from RCTs

RCTs are considered the gold standard for determining the effectiveness of a health intervention. This design is advantageous since assignment to the intervention group is randomized and if successful, ensures that all possible confounding variables (both measured and unmeasured), are equally distributed among the arms of the trial. By creating a control group with the same characteristics, RCTs can investigate the relationship between an intervention and outcome in an un-confounded manner, enabling causal inference.

RCTs are expensive and consequently, there are only five of trials investigating the efficacy of vitamin D supplementation for treatment of MS.(45–49) In general evidence is conflicting. For instance, one clinical trial involving 23 participants with relapse-remitting MS found no benefit of high-dose (13,000 IU) relative to lose-dose (1,000 IU)

vitamin D2 in terms of protection from new lesions(46). Conversely, a slightly larger study (N=66) investigating the effect of a weekly dose of 20,000 IU of vitamin D3 found that treatment significantly decreased the development of these same lesions.(47) Two additional trials detected no difference in MS disability measures or relapse rate between treatment and control arms.(48,49) The inconclusive results of these trials should be interpreted with caution. Considering the small sample and short follow-up period (**Table 1**), too few participants experienced the primary endpoints to allow for robust statistical inferences.

In an attempt to improve power, a 2013 meta-analysis by James et al. combined these 5 studies to explore the relationship between vitamin D supplementation and MS relapse rate (**Table 1**).(2) This meta-analysis including a total of 129 participants treated with high-dose vitamin D and 125 controls, again yielded an inconclusive finding (OR=0.98, 95% CI=0.44-2.17) (**Table 1**).(2) Another systematic-review analyzing at these same 5 trials, elected to not perform a meta-analysis stating that the trials were too heterogeneous to combine.(50) Since these 5 trials investigated different endpoints with varying control groups, dosages of vitamin D, and quality this may be a more reasonable approach as per PRISMA guidelines.(51)

Thus there is a lack of high-quality evidence from RCTs in order to provide conclusive evidence of vitamin D's causal role in MS etiology. While there are more trials ongoing,(2,52) the main problem resides in the small sample size of the previous and forth-coming trials. As such, these small short-term trials have inadequate statistical power to detect a realistic treatment effect of vitamin D. Furthermore, all trials published to-date investigated supplement dosages far exceeding the daily allowance

recommended by the Institute of Medicine (IOM) in their 2011 review.(53) Therefore the effect of vitamin D supplements falling within the physiological range remains untested for MS. Additionally, these trials (both previous and forthcoming) have investigated the use of vitamin D for the treatment of MS, as measured by varying disease activity indicators (i.e. the appearance of new lesions or change in EDSS score). However results from observational studies suggest that vitamin D may play role in MS incidence. Thus the role of vitamin D in the prevention of MS is unknown.

The Feasibility of Applying MR Methods to Vitamin D and MS

MR offers an alternative approach that can provide evidence to support, or refute, the causality of vitamin D in MS etiology. The implementation of the MR design in the context of vitamin D and MS is feasible. First, vitamin D is a modifiable risk factor with a strong genetic contribution as demonstrated by a 2010 GWAS (N= 33,996) investigating the genetic determinants of vitamin D insufficiency among individuals of European descent.(54) This study measured 25OHD across 15 European and North American cohorts and identified four SNPs that achieved genome-wide significance for 25OHD. Notably all four of these SNPs map to genes involved in vitamin D synthesis, transport and metabolism,(54) which is an advantage for the purposes of MR since an associated vitamin D mechanism decreases the possibility of pleiotropy. Moreover since these SNPs combined to explain 2.5% to 4% of the variance in 25OHD with effect sizes similar to that of vitamin D supplementation and season (**Fig. 1**),(54) MR analysis will be adequately powered to make robust inferences. Lastly, their validity as instruments has also previously been tested, with the study concluding that the four SNPs satisfied all

necessary MR criteria (more details provided in **Chapter 3**)(55). The second component necessary to complete this MR are the corresponding effect estimates for these SNPs upon MS. These can be obtained from International Multiple Sclerosis Genetics Consortium (IMSGC), which has employed GWAS to interrogate the genetic susceptibility of MS (in upwards of 14,498 cases and 24,091 healthy controls).(56,57) Despite the availability of necessary data, this study question has yet to be examined by MR methods. Thus we elected to perform an MR analysis to determine whether genetically decreased vitamin D influences risk of MS.

Table 1:

Study (author, date)	Sample size	Vitamin D measurement /dosage	Main outcome	Number of MS cases	Follow-up time	Summary of results
Observational studies *Cohort*						
Munger et al. 2004 (1)	NHS I = 92,253 NHS II = 95,310	Self-reported vitamin D intake (≥ 400 UI/day)	Incident MS	173	NHS I = 20 years NHS II = 10 years	Decrease in MS incidence for women taking ≥ 400 IU daily
Simpson et al. 2010 (36)	145	25OHD	MS relapse rate	145	3 years	Decrease in relapse rate per 10 nmol/L increase in 25OHD
Runia et al. 2011 (35)	73	25OHD	MS relapse rate	73	~1.7 years	Decrease in relapse rate per doubling of 25OHD
Munger et al. 2011 (30)	NHS I = 73,938 NHS II = 45,848	Self-reported adolescent vitamin D intake (≥ 400 UI/day)	Incident MS	379	NHS I = 28 years NHS II = 16 years	Suggestive, but non-significant, decrease in MS incidence for women ≥ 400 IU daily

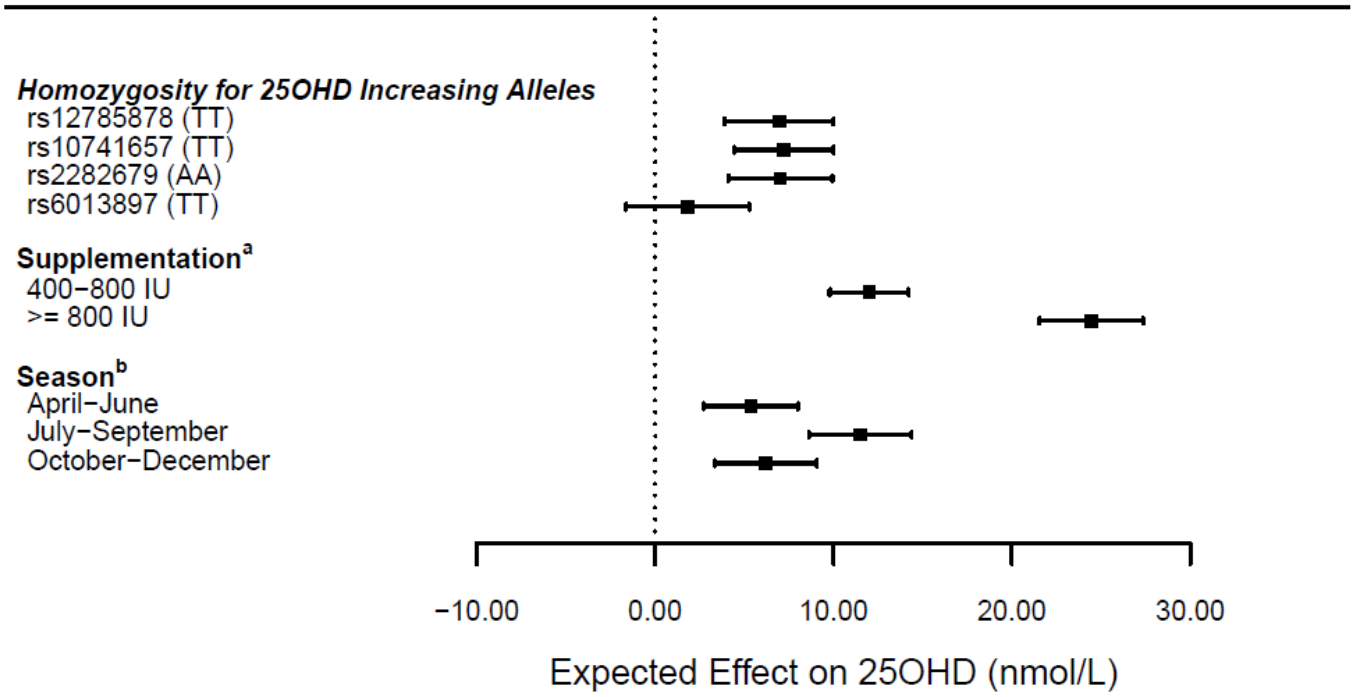
Mowry et al. 2012 (34)	469	25OHD	New lesions, relapse rate, disability	469	5 years	Decrease in EDSS score, new lesions per 10ng/mL increase in 25OHD
Ascherio et al. 2014 (15)	465	25OHD	New lesions, relapse rate, disability	465	5 years	Decrease in new lesions per 50 nmol/L increase in 25OHD Suggestive but non-significant decrease in relapse rate Lower annualized change in EDSS score observed for 25OHD level ≥ 50 nmol/L
Case-Control						
Munger et al. 2006 (17)	444	25OHD	Incident MS	148	12 years	Decrease in odds of MS per 50 nmol/L increase in 25OHD

Kragt et al. 2009 (31)	213	25OHD $1,25(OH)_2D$	Incident MS	103	6 months	Decrease in odds of MS per 10 nmol/L increase in 25OHD
RCTs						
Burton et al. 2010 (45)	47	Treatment= escalating doses of vitamin D3 (up to 40,000 IU/day (28 weeks), followed by10,000 IU/day (12 weeks) Control= ≤ 4,000 IU if desired	Tolerability of high-dose vitamin D3 Relapse rate, disability	47	1 year	Study did not calculate measures of effect Favorable trend observed in treatment group
Kampman et al. 2012 (49)	68	Treatment= 20,000 IU/week vitamin D3 (c Control= continuation of baseline supplements dosage	Relapse rate, disability	68	~ 2 year	No significant differences in relapse rate and disability measures between trial arms
Shayganeejad et al. 2012 (48)	50	Treatment= 0.25 µg/L 1,25-dihydroxyvitamin D3 per day (2 weeks), 0.5 µg/L	Relapse rate, disability	50	1 year	No significant differences in change in EDSS score and relapse

		per day (51 weeks) Control= placebo				rate between arms of trial
Soilu-Hanninen et al. 2012 (47)	66	Treatment= 20,000 IU/week vitamin D3 Control= placebo	T2 burden of disease, new lesions, relapse, disability	66	1 year	Significant decrease in T1 lesions Suggestive but non-significant in improvement disability indicators
Stein et al. 2012 (46)	23	Treatment= 13,00 IU vitamin D2 Control=1,000 IU	New lesions, relapse, disability	23	6 months	No difference in new lesions and relapse rate between arms of trials Suggestive, but non-significant improvement in EDSS score
Meta-analysis						
James et al. 2013 (2)	254	Treatment= variable dosages and forms of vitamin D	Relapse	125	26 – 96 weeks	Inconclusive effect of vitamin D upon relapse rate

		Control= placebo or low-dose vitamin D				

Fig. 1: Forest Plot of the SUNLIGHT SNPs, Supplementation and Season on 25OHD in the CaMos Cohort



Boxes and error bars represents the estimate and its 95% CI from a linear regression model of 25OHD (nmol/L) in the CaMos population. The regression model included the 4 SUNLIGHT SNPs, BMI, age, age^2, sex, supplementation status (coded as <400 IU, 400-800 IU, >= 800 IU and season of measurement (coded as for January-March, April-June, July-September, October-December). The estimated effect of homozygosity for the 25OHD increasing alleles is comparable to the estimated effects of supplementation and season. β coefficients of the regression model are provided in **Appendix**.

[a] represents the effect of supplementation relative < 400 IU daily

[b] represents the effect of season relative to blood measurement taken in January, February or March

Chapter 3 – Vitamin D and Risk of Multiple Sclerosis: a Mendelian randomization study

Lauren E Mokry,[1] Stephanie Ross,[1] Omar S. Ahmad,[1,2] Vincenzo Forgetta,[1,2] George Davey Smith,[4] Aaron Leong,[5] Celia MT Greenwood,[6] George Thanassoulis,[2,3] J. Brent Richards(*)[1,2,3,7]

Author Affiliations

[1]Centre for Clinical Epidemiology, Department of Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, McGill University, Montreal, Quebec, Canada

[2]Department of Medicine, McGill University Montreal, Quebec, Canada

[3]Department of Human Genetics, McGill University, Montréal, Québec, Canada

[4]MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK

[5]Division of General Internal Medicine, Massachusetts General Hospital and Department of Medicine, Harvard Medical School, Boston, Massachusetts, United States

[6]Department of Oncology, Epidemiology, Biostatistics and Occupational Health, and Human Genetics, McGill University, Montreal, Quebec, Canada

[7]Department of Twin Research and Genetic Epidemiology, King's College London, United Kingdom

Manuscript as published in PLOS Medicine.

Citation: Mokry LE, Ross S, Ahmad OS, Forgetta V, Smith GD, et al. (2015) Vitamin D and Risk of Multiple Sclerosis: A Mendelian Randomization Study. PLOS Med 12(8): e1001866. doi: 10.1371/journal.pmed.1001866

ABSTRACT

Background: Observational studies have demonstrated an association between decreased vitamin D levels and risk of multiple sclerosis (MS), however it remains unclear whether this relationship is causal. We undertook a Mendelian randomization (MR) study to evaluate whether genetically lowered vitamin D levels influence risk of MS.

Methods and Findings: We identified single nucleotide polymorphisms (SNPs) associated with 25-hydroxyvitamin D (25OHD) levels from the SUNLIGHT consortium, the largest (n = 33,996) genome-wide association study to date for vitamin D. Four SNPs were genome-wide significant for 25OHD levels (P-values ranging from $6x10^{-10}$ to $2x10^{-109}$) and all SNPs lay in, or near, genes strongly implicated in separate mechanisms influencing 25OHD. We then ascertained their effect on 25OHD levels in 2,347 participants from a population-based cohort, the Canadian Multicentre Osteoporosis Study and tested the extent to which the 25OHD-decreasing alleles explained variation in 25OHD levels. We found that the count of 25OHD-decreasing alleles across these four SNPs was strongly associated with lower 25OHD levels (n = 2,347, F-test statistic = 49.7, P = $2.4x10^{-12}$). Next, we conducted an MR study to describe the effect of genetically lowered 25OHD on the odds of MS in the International MS Genetics Consortium, the largest genetic association study to date for MS (including up to 14,498 cases and 24,091 healthy controls). Alleles were weighted by their relative effect on 25OHD levels and sensitivity analyses were performed testing MR assumptions. MR analyses found that each genetically determined standard deviation

decrease in log transformed 25OHD levels conferred a 2.0-fold increase in odds of MS (95% CI: 1.7-2.5; P= $7.7x10^{-12}$; I^2 = 63%, 95% CI: 0%-88%) . These results persisted after sensitivity analyses excluding SNPs possibly influenced by population stratification or pleiotropy (OR = 1.7, 95% CI: 1.3-2.2; P = $2.3x10^{-5}$; I^2 = 47% 95% CI: 0%-85%) and including only SNPs involved in 25OHD synthesis or metabolism ($OR_{synthesis}$= 2.1, 95% CI: 1.6-2.6; P=$1x10^{-9}$ and $OR_{metabolism}$= 1.9, 95% CI: 1.3-2.7; P=0.002). While these sensitivity analyses decrease the possibility that pleiotropy may have biased the results, residual pleiotropy is difficult to exclude entirely.

Conclusions: Genetically lowered 25OHD levels are strongly associated with increased susceptibility to MS. Whether vitamin D sufficiency can delay, or prevent, MS onset merits further investigation in long-term randomized controlled trials.

INTRODUCTION

Multiple sclerosis (MS) is the most common permanent neurological disorder affecting young adults.(58) It is a debilitating autoimmune condition that presents early in life with a mean age of onset of 28-31 years. Epidemiological studies have indicated that the prevalence of MS varies geographically, such that regions of higher latitude and decreased levels of sunlight exposure have a higher prevalence of MS.(20,59) Since circulating levels of vitamin D, as measured by 25-hydroxyvitamin D levels (25OHD, the clinical determinant of vitamin D status), are partially derived from sunlight exposure, it has been suggested that 25OHD deficiency may be the causal risk factor mediating this latitudinal gradient.(60) Further evidence to support the vitamin D hypothesis arose from the Nurse's Health Study, which reported a protective effect on MS for women who had high levels of daily vitamin D intake.(1) Lower vitamin D levels have also been associated with higher rates of MS relapses(61) and higher MS-specific disease activity and disability.(16) Vitamin D has important effects upon the immune system, and its immune-modulating effects have been observed in multiple cell-culture experiments,(62) providing possible biologic mechanisms whereby vitamin D may influence MS risk.

To date, there has been one published meta-analysis that investigated the effect of vitamin D supplementation on MS relapse in five randomized controlled trials (RCTs) among 254 individuals.(2) The authors reported that the effect of high dose vitamin-D treatment on MS relapse was inconclusive (OR: 0.98, 95% CI 0.45–2.16), and that these trials had important methodologic limitations, such as small sample sizes and

short duration of vitamin D treatment. In contrast, two other non-blinded trials demonstrated improved clinical outcomes with vitamin D therapy; however, disease activity, and/or MRI changes were not the primary outcome of these trials.(45,63) Importantly, these trials test whether vitamin D can *treat* MS, but provide no insight as to whether vitamin D can *prevent* MS.

Consequently, clinical practice guidelines for the treatment of MS(64) do not include vitamin D therapy. This is at least partially attributed to the possibility of confounding in the above observational studies. Additionally, observational studies are also prone to reverse causation; where, for example, individuals with MS may spend less time outdoors and as a result have lower circulating 25OHD levels. However, if decreased 25OHD levels were causally associated with MS, this could have important implications since vitamin D insufficiency, as defined as 25OHD levels < 50 nmol/L, is common and increasing in prevalence. This was observed in the National Health and Nutrition Examination Survey (NHANES) where in 2005, 41.6% of adult Americans were found to be vitamin D insufficient with mean 25OHD levels decreasing from 75 nmol/L in NHANES in 1988 to 50 nmol/L in 2006.(5,65)

In the absence of high-quality RCT data, the principles of Mendelian randomization (MR) can be applied to strengthen or refute the causality of biomarkers in disease etiology.(66) MR analysis uses genetic associations to test the effects of biomarkers, such as 25OHD, on the risk of disease. This approach, which is conceptually similar to an RCT, is based on the principle that genetic variants are randomly allocated at

meiosis and consequently these genetic variants are independent of many factors that bias observational studies, such as confounding and reverse causation. MR methods have been used previously to question the role of HDL(9) and CRP(67) in predisposition to cardiovascular disease, and have provided strong evidence that PCSK9 inhibition prevents cardiovascular disease.(68) MR methods may be of particular relevance to understanding the etiology of MS since date of disease onset is often poorly recognized clinically and MR studies assess the effect of life-time exposures.

Here we adopted an MR design to clarify whether 25OHD levels lie in the causal pathway for MS susceptibility. In order to assess whether reduced levels of 25OHD are associated with an increased risk of MS, we selected genome-wide significant single nucleotide polymorphisms (SNPs) as identified by the SUNLIGHT consortium, the largest genome-wide association study (GWAS) published to date for 25OHD levels. Next, we estimated the effect of each of these SNPs upon 25OHD levels in the Canadian Multicentre Osteoporosis Study (CaMos) and tested their validity as instrumental variables for MR analyses. Finally, we applied the principles of MR to provide evidence of the association of a lifetime of genetically lowered 25OHD levels on MS risk using data from the Multiple Sclerosis Genetics Consortium (IMGSC).

METHODS

SNP selection and Data Sources

Genetic variants associated with 25OHD levels at a genome-wide significant level ($p < 5x10^{-8}$) were obtained from the SUNLIGHT Consortium(54), a genome-wide association (GWAS) study consisting of 33,996 individuals of European descent from 15 cohorts. 25OHD levels in this study were measured either by radioimmunoassay, chemiluminescent assay, ELISA or mass spectrometry. Given that different cohorts used different methods to measure 25OHD levels, results were combined across cohorts in the SUNLIGHT Consortium using Z-score-weighted meta-analysis.

The Canadian Multicentre Osteoporosis Study (CaMos) was used to estimate the effect of each genome-wide significant SNP on 25OHD levels, since effect of each SNP upon 25OHD levels could not be used from the SUNLIGHT Consortium, due to the Z-score meta-analytic approach employed.(69) CaMos is a large population-based cohort, and was amongst the largest included in the replication phase of the SUNLIGHT consortium. It includes 2,347 individuals who were genotyped using TaqMan genotyping at the same genome-wide significant vitamin D loci found in the SUNLIGHT consortium.

To obtain precise estimates for the association of 25OHD on MS we tested the effect of each genome-wide significant SNP for vitamin D levels in the Multiple Sclerosis Genetics Consortium (IMGSC) Immunochip study, the largest international genetic consortium for Multiple Sclerosis involving 14,498 MS cases and 24,091 healthy controls.(57) All subjects were of European ancestry and were genotyped using the

Immunochip array, which is a custom array designed to interrogate SNPs with potential immune system effects. Cases were defined as individuals diagnosed by a neurologist according to recognized diagnostic criteria dependent on laboratory and clinical information.(70–72) When data were not available for a specific SNP in the IMGSC Immunochip study, we used data from the second largest MS genetic study, the IMSGC and Wellcome Trust Case Control Consortium 2 (WTCCC2) study, which included 9,772 cases and 6,332 controls taken from the WTCCC2 common control set.(56)

SNP Validation and Effect Sizes

Linkage disequilibrium Assessment

One assumption of MR studies is that the selected SNPs must not be in linkage disequilibrium (LD) since if a selected SNP is highly correlated with other risk factor loci, this may result in confounding.(66) In order to verify this, we measured LD between all selected SNPs using CEU samples from the 1000 Genomes Project (n=94).(73)

Pleiotropy Assessment

MR analyses assume that the chosen SNPs do not exert pleiotropic effects on the outcome (in this case, MS) by operating through biological pathways independent of the exposure (in this case, 25OHD levels). However, in MR, a SNP may influence the outcome via other factors, if the SNP acts upon the other factors *through* the exposure itself.(74)Previous work has assessed possible pleiotropic actions of the 25OHD-related SNPs used in our analysis by investigating the association between 25OHD-related SNPs and clinical traits in the 1958 British Birth Cohort which included 6,877

participants of European descent.(55) In this cohort, no associations were found between these SNPs and relevant potential pleiotropic pathways, such as sun exposure, time outside, physical activity, oily fish consumption, smoking, alcohol consumption, body mass index (BMI), abdominal obesity and social class ($P > 0.05$ for all).(55) However, we note that some of these factors, such as sun exposure, time outside, BMI and abdominal obesity could act at least partially through the vitamin D pathway. Furthermore, SNPs associated with 25OHD levels did not associate with other biomarkers (including C-reactive protein, IgE levels, von Willebrand's factor, tissue plasminogen activator, D-dimer, fibrinogen, triglycerides, HDL, LDL or total cholesterol, forced expiratory volume, diastolic blood pressure, IGF-1, HbA1c) and no interactions were observed between the SNPs, biomarkers and 25OHD levels. Additional details are provided in **S1 Table**.

To further explore sources of pleiotropy, we also conducted a systematic literature search of gene name, gene mutation and protein name to examine the published literature for possible pleiotropic mechanisms for any of our selected SNPs on MS and autoimmunity using PubMed. Details of this method are described in the **S1 Methods**.

Population Stratification Assessment

The 1958 British Birth Cohort has previously assessed the potential for population stratification of the 25OHD-associated SNPs, which is a potential source of bias in MR studies since differences in minor allele frequencies between populations may cause the SNP to be associated with both the ancestry and the outcome.(55) In the 1958 Birth

Cohort, each SNP was tested for association with geographic region, which was dichotomized as South and Middle UK (South East, South West, and Greater London, East Anglia, Midlands, and Wales) vs. Northern England (North, North West, and Yorkshire and the Humber) and Scotland. We then verified potential population stratification of each SNP by testing their association with self-declared ethnicity in the CaMos cohort and finally tested the association of each SNP with non-European status, defined as exclusion from the European cluster in principal component analyses (PCA).

Effect Size Estimates of SUNLIGHT SNPs upon 25OHD Level

To obtain the effect of each SNP upon 25OHD levels required for our MR analysis, we tested the additive effect of each minor allele on natural log-transformed 25OHD levels in the CaMos study, while controlling for sex, age, age squared, BMI and season of 25OHD measurement (using categorical variables for summer [July-September], autumn [October-December], winter [January-March] and spring [April-June]).(54) [Rationale for this adjustment set and results using 25OHD on the absolute scale are provided in the **Appendices**]. Ethnicity was checked by self-report and verified using PCA analyses. Individuals that did not cluster with other Europeans were excluded from this analysis and were not used to measure the effect of each SNP upon 25OHD to prevent population stratification from confounding our results. A count of 25OHD decreasing alleles was calculated for each subject in the CaMos cohort. This allele count was tested for an association with natural log-transformed 25OHD levels using linear regression, which had been residualized for the above covariates and the F-statistic for the allele score was reported. The multiply-adjusted natural log-transformed

25OHD levels were then assessed for each category of allele count and a non-parametric trend test across these allele counts was computed.

Association of SUNLIGHT SNPs with MS Susceptibility

In order to increase study power and obtain the most precise estimates of the association of 25OHD-associated SNPs upon risk of MS, we used summary-level data from the IMSGC Immunochip Study, if available (as described above). However, the IMSGC Immunochip genotyping array used is not genome-wide, so not all SNPs were captured in this experiment. If a SNP was not included in the IMSGC Immunochip study, then summary statistics from the second largest genotyped cohort, the IMSGC/WTCCC2, were selected. In the event that a SNP was not genotyped in either cohort, summary statistics for a perfect proxy SNP, defined as a surrogate SNP with perfect LD (r^2 = 1.0) to the SNP interest, was selected. LD for proxy SNPs was calculated using CEU samples from the 1000 Genomes Project (n=94) since the IMSGC samples are of the same ancestry.(73) We then assessed whether each SNP was associated with risk of MS, applying a Bonferroni correction, where statistical significance was declared at P ≤0.05/*n* where *n* is the number of SNPs associated with 25OHD levels from the SUNLIGHT consortium.

Mendelian Randomization Estimates

We conducted our MR analysis by assessing the effects of the SNPs upon risk of MS, weighting the effect of each SNP by the magnitude of its effect upon 25OHD levels. In this study design, which has been described previously,(8,75,76) the independent SNPs

evaluate the association of exposure to genetically lowered 25OHD with MS risk. These individual estimates were then pooled using statistically efficient estimators formally analogous to those of inverse-variance weighted meta-analysis.(77) We carried out a meta-analysis of estimates obtained from individual 25OHD decreasing alleles using both fixed-effects and random-effects models to obtain pooled estimates for the combined effect of the 25OHD SNPs on MS.

Specifically, let x and y denote the centered and scaled natural log 25OHD and log-odds MS traits, respectively, and suppose these are related by the linear structural equation: $y = \alpha x + \eta$. Here, η is a stochastic error term, and in general x and η are correlated because of confounding. The parameter α quantifies the causal effect of x on y, and is thus the parameter we seek to estimate. Let u_i denote the allele dosage variable of the i^{th} genetic variant. Let γ_i and β_i denote effect-size estimates (derived from GWAS data) of u_i on the exposure (change in natural log 25OHD levels) x and outcome (change in log odds of MS) y, respectively, and let $s(\beta_i)$ denote the standard error of β_i. Then the MR estimate associated with the i^{th} genetic variant is $\alpha_i = \beta_i / \gamma_i$, and the variance of this estimate is $v_i = (s(\beta_i) / \gamma_i)^2$. Define the precision of the i^{th} MR estimate of α by $w_i = 1/v_i$. The inverse-variance-weighted fixed-effects estimate is then $\alpha_{fixed} = \sum_{i=1}^{n} w_i\alpha_i / \sum_{i=1}^{n} w_i$, and the standard error $s(\alpha_{fixed})$ of this estimate is given by $s(\alpha_{fixed}) = (\sum_{i=1}^{n} w_i)^{-1/2}$. We observe that α_{fixed} may also be interpreted as the regression coefficient resulting from the generalized linear regression of the outcome effect sizes β_i on the

exposure effect sizes γ_i assuming heteroskedastic errors; in this regression, the i^{th} error term has a variance equal to $s(\beta_i)^2$, and the offset coefficient in the regression is zero..

The random-effects estimate α_{random} and its standard error $s(\alpha_{random})$ were also constructed from the individual estimates using standard methods,(78) in which the weights are adjusted to account for the intrinsic variability (or *heterogeneity*) in the effect size. Heterogeneity may be quantified with the parameter I^2, which reports the fraction of the total variance in the meta-analytic estimate that is due to intrinsic variability in the effect-size, as distinct from variability arising due to measurement error.(79) The random-effects estimate α_{random} and its standard error $s(\alpha_{random})$ are given by equations analogous to those for α_{fixed} and $s(\alpha_{fixed})$, in which the weights assigned to individual estimates are adjusted to take into account heterogeneity in the effect-size.

For all MR meta-analyses, we report estimates using both fixed-effects models and random effects model. The effect-sizes for each meta-analysis is reported in the main results as the effect of a standard deviation (SD) change in natural log-transformed 25OHD levels, since this metric is more interpretable than an arbitrary difference. This measure is given by $exp(\alpha_{fixed})$ for the fixed-effects model, and by $exp(\alpha_{random})$ for the random-effects model. We also report the I^2 as an assessment of heterogeneity.

In order to provide a better clinical interpretation of a one SD change in natural log transformed 25OHD levels, we selected three clinically relevant 25OHD thresholds for vitamin D status (25 nmol/L for vitamin D deficiency, 50 nmol/L for vitamin D

insufficiency and 75 nmol/L for vitamin D sufficiency).(53) These values were converted to the natural log scale because the magnitude of a one SD change is not constant on the untransformed scale. For each of these natural log transformed 25OHD levels, we then calculated a one SD increase in natural log transformed 25OHD. To obtain 25OHD levels which correspond to circulating levels in units of nmol/L, we then back-transformed these values.

Sensitivity Analyses

MR estimates were re-calculated after exclusion of SNPs potentially influenced by pleiotropy or population stratification. Since SNPs associated with 25OHD levels in the SUNLIGHT consortium influence either 25OHD synthesis or 25OHD metabolism,(55) we elected to perform a stratified MR analysis where SNPs involved in either 25OHD synthesis or metabolism were analyzed separately.

RESULTS

SNP Selection and Validation

SNP selection

A schematic representation of the MR study design is presented in **Fig. 2**. The SUNLIGHT Consortium identified four SNPs as genome-wide significant for 25OHD levels.(54) These included: rs2282679 in *GC* (association with 25OHD $P = 1.9x10^{-109}$), rs12785878 near *DHCR7* ($P = 2.1x10^{-27}$), rs10741657 near *CYP2R1* ($P = 3.3x10^{-20}$) and rs6013897 in *CYP24A1* ($P = 6.0\times10^{-10}$). We selected these SNPs for our MR study since all were strongly associated with 25OHD levels and map to genes implicated in the modulation of 25OHD levels through distinct mechanisms.(80) Specifically, *GC* encodes the vitamin D binding protein (DBP), a group-specific component of serum globulin. DBP acts as the principal protein carrier for 25OHD, transporting 80-90% of 25OHD to target organs.(81–83) The *DHCR7* gene product is known to convert 7-dehydrocholesterol to cholesterol, providing a substrate for vitamin D production. CYP2R1 is a regulator of 25OHD synthesis through 25-hydroxylation of vitamin D in the liver, the first activation step,(84) and lastly, CYP24A1 inactivates $1\alpha25(OH)_2D$ rendering inactivation of the active form of vitamin D (**Fig. 3**). Therefore, all SNPs used in this study map near genes strongly implicated in vitamin D synthesis, transport or metabolism. Notably, all 4 SNPs lie in intergenic or intronic regions, and presently the exact effect of each SNP on these enzymes is unknown. Nevertheless, all SNPs reside near genes strongly implicated in vitamin D synthesis or metabolism.(80)

Linkage Disequilibrium and Pleiotropy Assessment

There was no evidence of LD between any of these SNPs (all pairwise $r^2 \leq 0.01$) in the 1000 Genomes Project CEU samples. We note that only two of our SNPs, rs10741657 and rs12785878 were located on the same chromosome, which greatly decreases risk of confounding by LD. As described above, none of the four 25OHD SNPs was associated with relevant pleiotropic pathways in the 1958 British Birth Cohort.

Undertaking a literature review for possible pleiotropic pathways, we found no evidence for pleiotropic mechanisms for the vitamin-D metabolism SNPs: rs10741657 (*CYP2R1)* and rs6013897 (*CYP24A1)*. For rs2282679 (*GC*), we found that its encoded protein, DBP has been associated with macrophage activation and may modulate T-cell response to vitamin D.(85) Elevated DBP levels are also found in the cerebrospinal fluid of patients with Alzheimer's disease(86) and MS,(87) and have been linked to the progression of MS in rats.(88) It has been argued that DBP can act independently of vitamin D to produce clinical phenotypes therefore we undertook sensitivity analyses excluding the rs2282679 (*GC*) in our MR analyses. Genetic variation in *DHCR7* appears to cause Smith-Lemli-Opitz syndrome, a clinical phenotype relating to cholesterol deficiency. Given that a recent study suggested an inter-dependence of cholesterol and vitamin D pathways in the etiology of MS,(89) we queried the association of rs12785878, in the largest publically available GWAS consortium results for lipids, the Global Lipids Genetics Consortium,(90) and found that this SNP was associated with a minimum p-value of 0.043 across all lipid traits, suggesting that the SNP is not strongly associated with cholesterol.

Population Stratification Assessment

Previous reports from the 1958 British Birth Cohort demonstrated that rs12785878 (*DHCR7*) was associated with geographic region.(55) Since rs12785878 is unevenly distributed across geography and the prevalence of MS varies by geographic location, a potential surrogate for local ancestry,(60) we tested to whether this SNP was associated with non-European status in CaMos using PCA. The SNP rs12785878 was strongly associated with non-European status in the CaMos cohort ($P = 2.7x10^{-13}$). No other SNP showed any evidence of correlation with non-European status (P >0.5 for all other SNPs). Given this possible relationship with population stratification, we undertook MR sensitivity analyses excluding the rs12785878 (*DHCR7*) variant.

Association of SUNLIGHT SNPs with 25OHD Levels

Table 2 displays the four SNPs that achieved genome-wide significance for 25OHD levels in the SUNLIGHT consortium, and describes their association with 25OHD.(54) Each of these SNPs explained an important proportion of the population-level variance in 25OHD levels, as reflected by the F-statistics. The count of 25OHD decreasing alleles across these four SNPs was strongly associated lower 25OHD levels in the CaMos population, residualized for age, season, sex and BMI (F-statistic = 49.7, r^2 = 2.44%, P for allelic score = $2.4x10^{-12}$). **Fig. 4** shows the mean of 25OHD levels for individuals with increasing counts of 25OHD decreasing alleles (non-parametric trend test P = $3.3x10^{-19}$).

Table 2: Characteristics of SNPs used as instrumental variables

Vitamin D Results								MS Results		
Locus	**Chr**	**25OHD Associated SNP**	**25OHD Decreasing Allele**	**Allele Frequency**	**Effect on 25OHD***	**P-value for Association with 25OHD#**	**F-Statistic for 25OHD♮**	**OR (95% CI) for MS**	**P-Value for Association with MS**	**Study**
CYP2R1	11	rs10741657	C	0.62	-0.052	$3.3x10^{-20}$	18.78	1.05 (1.02-1.09)	$3.9x10^{-3}$	Immunochip (57)
DHCR7	11	rs12785878	G	0.27	-0.056	$2.1x10^{-27}$	18.29	1.11 (1.07-1.15)†	$8.7x10^{-9}$	Immunochip (57)
GC	4	rs2282679	C	0.30	-0.047	$1.9x10^{-109}$	13.38	1.04 (1-1.08)	$6.2x10^{-2}$	WTCCC2 (56)
CYP24A1	20	rs6013897	A	0.19	-0.027	$6.0x10^{-10}$	3.13	1.07 (1.03-1.11)‡	$1.7x10^{-3}$	WTCCC2 (56)

*Effect on Multiply Adjusted Natural Log-Transformed 25OHD levels in the CaMos Cohort
#P-values derived from the SUNLIGHT Consortium
♮F-Statistic derived from Multiply-Adjusted Natural Log-Transformed 25OHD levels in the CaMos Cohort
†SNP rs12785878 was not available for MS data. Therefore SNP rs4944958 was used as a proxy (r^2 between these two SNPs = 1.0)
‡SNP rs6013897 was not available for MS data. Therefore SNP rs17217119 was used as a proxy (r^2 between these two SNPs = 1.0)

Association of SUNLIGHT SNPs with MS Susceptibility

Summary statistics for two of the four 25OHD-associated SNPs, (rs10741657 at *CYP2R1* and rs12785878 at *DHCR7)* and their association with MS was taken from the IMSGC Immunochip study (**Table 2**). rs12785878 at *DHCR7* was not directly genotyped in the Immunochip study, however, a perfect proxy for rs12785878, rs4944958, was used (r^2 = 1.0 between rs12785878 and rs4944958 in the 1000 Genomes Project CEU samples). Summary statistics for the remaining two SNPs, (rs6013897 at *CYP24A1* and rs2282679 at *GC*), were taken from the second largest MS genetic association study: IMSGC/WTCCC2. SNP rs6013897 at *CYP24A1* was not present in the IMSGC/WTCCC2 dataset and therefore a perfect proxy SNP for rs6013897, rs17217119, was used (r^2 = 1.0 between rs17217119 and rs6013897 from the 1000 Genomes Project CEU samples).

All four 25OHD-decreasing alleles associated with an increased risk of MS (**Table 2**). rs12785878 (*DHCR7),* achieved genome-wide significance for MS risk while two 25OHD-decreasing alleles (rs10741657 and rs6013897) were moderately associated with MS risk ($P=3.9x10^{-3}$ and $P=1.7x10^{-3}$, respectively). The 25OHD-decreasing allele rs2282679[C] (*GC)* was not significantly associated with MS risk (P=0.062) (**Table 2**). However, three of the 25OHD associated SNPs (rs12785878, rs10741657 and rs6013897) remained associated with MS, after a Bonferroni correction for the number of independent SNPs ($P \leq 0.05/4 = 0.0125$).

Mendelian Randomization Analysis for the Association of 25OHD with MS Risk

In order to estimate the association of genetically lowered 25OHD upon MS, we used a fixed-effects model in which all four 25OHD-decreasing alleles of the MR set were included. We observed that each standard deviation decrease in natural log-transformed 25OHD levels was associated with an increased risk of MS (OR = 2.02, 95% CI: 1.65-2.46; P= 7.72×10^{-12}) (**Table 3** and **Fig. 5**). Given that the I^2 estimate of heterogeneity was somewhat increased (I^2 = 63%, 95% CI:0%-88%), we also undertook random effects meta-analysis, which generated similar findings (OR= 2.07, 95% CI: 1.45-2.96; P=5.7×10^{-5}) (**Table 3** and **S1 Fig.**)). We note that since our model included only 4 SNPs, the 95% CIs of the I^2 statistic are wide and consequently heterogeneity cannot be accurately measured using this parameter. In addition, due to potential effects of population stratification and pleiotropy, we undertook a sensitivity analysis by excluding the rs12785878 SNP (*DHCR7)*. Despite removal of this variant we observed clear association of genetically lowered 25OHD levels on the risk of MS (OR= 1.72, 95% CI: 1.34-2.21; P=2.28×10^{-5}; I^2 = 47.3% 95% CI: 0%-85%) (**Table 4** and **Fig. 6**), which remained significant using a random effects meta-analysis (OR= 1.82, 95% CI: 1.24-2.67; P=2.13×10^{-3}; I^2 = 47.3%) (**Table 4** and **S2 Fig.**). Removal of the rs2282679 (*GC*) that may possibly be influenced by pleiotropy did not influence the MR results using a fixed effects or random effects model (OR = 2.1, 95% CI: 1.7-2.7; P = 1.7×10^{-11}; I^2 = 67% and OR = 2.30, 95% CI: 1.5-3.6; P = 1.8×10^{-4}; I^2= 67%, respectively) (**S3 Fig. and S4 Fig.**) To further assess the effect of the independent vitamin D pathways on risk of MS, we analyzed SNPs near genes implicated in 25OHD synthesis (*DHCR7* and

CYP2R1) and metabolism (*GC*, and *CYP24A1*) separately and found that both strongly associated with increased risk of MS (**Table 5**) (**S5 Fig. and S6 Fig.**).

Table 3: MR estimate of the association of decreased 25OHD on the risk of MS

Model	OR (95% CI)*	P-Value	I^2 (95% CI)
Fixed Effects	2.02 (1.65-2.46)	7.72×10^{-12}	63 (0-88)
Random Effects	2.07 (1.45-2.96)	5.74×10^{-5}	63 (0-88)

*OR is expressed as the odds of MS for a one standard deviation decrease in natural log- transformed 25OHD levels

Table 4: MR estimate of the association of decreased 25OHD on the risk of MS, excluding the *DHCR7* locus.

Model	OR (95% CI)*	P-Value	I^2 (95% CI)
Fixed Effects	1.72 (1.34-2.21)	2.28×10^{-5}	47.3 (0-85)
Random Effects	1.82 (1.24-2.67)	2.13×10^{-3}	47.3 (0-85)

*OR is expressed as the odds of MS, for a one standard deviation decrease in natural log transformed 25OHD levels

Table 5: MR estimate of the association of decreased 25OHD on the risk of MS, stratified by SNPs near genes involved in 25OHD synthesis or metabolism using a fixed effects model

Model	OR (95% CI)*	P-Value
Synthesis	2.08 (1.64-2.63)	1.1×10^{-9}
Metabolism	1.86 (1.26-2.74)	1.7×10^{-3}

*OR is expressed as the odds of MS, for a one standard deviation decrease in natural log transformed 25OHD levels. Note that the 95% CI for the I^2 cannot be properly estimated given that there are only two SNPs per model.

The clinical equivalence of a one SD increase in natural-log 25OHD for vitamin D deficiency (25 nmol/L), vitamin D insufficiency (50 nmol/L) and vitamin D sufficiency (75 nmol/L) are shown in **Table 6**. We observed that for vitamin D deficient (25 nmol/L) individuals an increase in 25OHD levels to 36.9 nmol/L would be required to decrease the odds of MS by 50% while for vitamin D insufficient (50 nmol/L) and vitamin D sufficient (75 nmol/L) individuals an increase in 25OHD levels to 73.7 nmol/L and 110.6 nmol/L would similarly be required.

Table 6: Clinical equivalence of one SD natural log increase in 25OHD for various vitamin D thresholds

Clinically relevant 25OHD Level	25OHD Level Required to Decrease Odds of MS by 50%*
Vitamin D Deficient (25 nmol/L)	36.86 nmol/L
Vitamin D Insufficient (50 nmol/L)	73.72 nmol/L
Vitamin D Sufficient (75 nmol/L)	110.6 nmol/L

*****Expressed as the equivalent of a natural log transformed SD increase in 25OHD on the nmol/L scale

DISCUSSION

Using summary level data for MS and 25OHD levels from large European populations, our study demonstrated that a genetic decrease in natural log-transformed 25OHD by one standard deviation was associated with 2-fold increase risk of MS, providing strong evidence in support of a causal role of vitamin D in MS susceptibility. These findings were consistent with evidence from observational studies that have demonstrated that low vitamin D levels influence risk of MS and also reflect findings from functional studies which have implicated vitamin D as an important regulator in the expression of MHC class II genes.(91,92) This provides rationale to further investigate whether vitamin D supplementation may reduce MS susceptibility in those most at risk.

The identification of vitamin D as a causal susceptibility factor for MS may have important public health implications since vitamin D insufficiency is common,(5,65) and vitamin D supplementation is both relatively safe and cost-effective.(53) The importance of these findings may be magnified in high latitude countries, which have disproportionately higher rates of MS and also higher rates of vitamin D insufficiency.

A reasonable first step to understand the role of vitamin D therapy in delaying the onset or severity of MS would be to treat vitamin D insufficiency in those most at risk of developing MS. MS is often preceded by the clinically isolated syndrome, which is a first clinical episode compatible with MS, often accompanied by lesions on magnetic resonance imaging,[44] thereby providing a therapeutic window and rationale in which to intervene with vitamin D supplementation. On-going RCTs are currently assessing

the role of vitamin D supplementation for the treatment and prevention of MS(93,94) and may therefore provide needed insights into the role of vitamin D supplementation.

An important difference between MR and RCTs is that MR studies describe the association of a *lifetime of exposure* to vitamin D lowering alleles in the general population, whereas RCTs provide insights from supplementation for shorter periods in individuals at risk. Thus, long-term RCTs may be needed to adequately assess the impact of vitamin D supplementation in the prevention, or treatment of MS. Lastly, MR may be an ideal study design to understand risk factors for MS, given the long latency period between disease onset and diagnosis, since MR may permit the estimate of lifetime exposure to risk factors.

Our analysis has several strengths. First, by utilizing the random allocation of genetic variants, we were able to overcome potential confounding and reverse causation that may bias estimates from observational studies. Second, using data from the largest genetic consortia for 25OHD levels (n = 33,996) and MS risk (up to 14,498 cases and 24,091 controls) has enabled us to more precisely test our study hypothesis than if we had used individual-level data from a small study. Previous work has shown that the use of estimates from meta-analytic data for uncorrelated genetic variants are similarly efficient to individual-level data in MR studies.(8) Lastly, the findings from this study represent the association of a life-long exposure to reduced vitamin D levels in the general European population and in the absence of large-scale, long-term RCT data,

our findings provide a strong evidence in support of a causal role of low vitamin D levels in MS susceptibility.

Our study also has limitations. First, while we have provided evidence supporting a role for vitamin D in MS susceptibility, we cannot conclude that vitamin D plays a role in disease modulation after its onset. While MR is able to overcome the limitations which may bias observational studies, the possibility of residual pleiotropy could bias estimates in this study. However, in this study the main findings have remained robust to multiple sensitivity analyses, testing the pleiotropy assumption thereby decreasing the probability of bias due to pleiotropy. We also note that all four studied SNPs are located in or near 25OHD-associated genes and influence 25OHD levels through known and distinct mechanisms. Additionally the point estimate for each 25OHD decreasing allele, as well as the combined 25OHD synthesis and metabolism pathways, were independently associated with increased risk of MS. Therefore it is unlikely that pleiotropy strongly biased our results. Like most MR studies we cannot directly assess whether canalization, which is defined as compensatory feedback interactions, may have influenced our results.(7,13,66) However since canalization assumes that other physiologic mechanisms may attenuate the effect of genetically reduced 250HD levels, such feedback interactions would tend to bias results toward the null, In contrast, our study has generated results that are very distinct from the null.

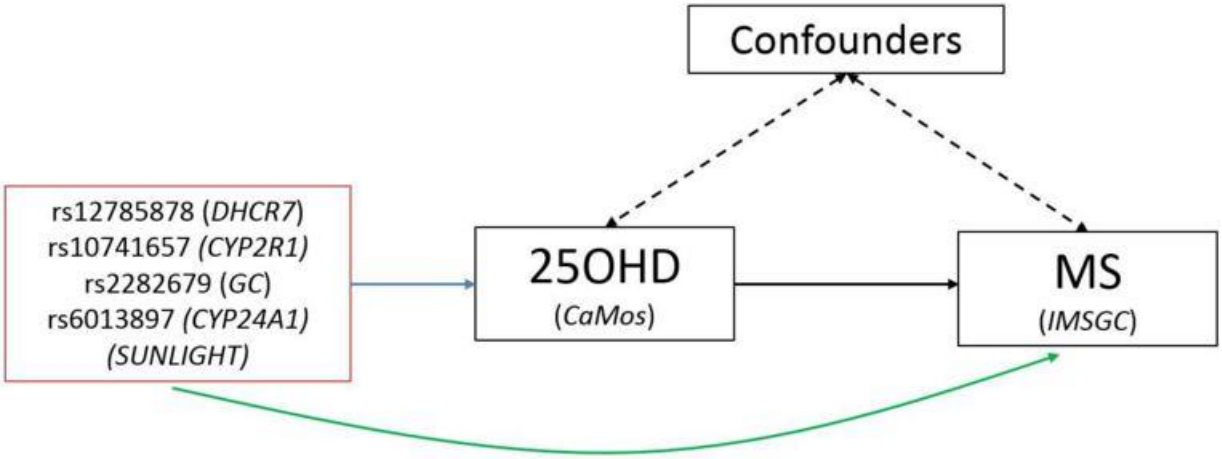
MR analyses using *DHRC7*, *GC*, *CYP24A1* and *CYP2R1* as instruments have been performed in the past.(42,95–98) We and others have recently provided evidence from

MR that low vitamin D levels do not increase insulin resistance(95) or the risk of type 2 diabetes,(95,96) and coronary heart disease,(95) but do increase the risk of type 1 diabetes(97) and possibly blood pressure.(98) Interestingly MR has shown that 25OHD levels are directly influenced by body mass index, and converse effects are likely to be small.(42) Thus while observational associations between 25OHD and two autoimmune conditions – type 1 diabetes and now MS – have been supported by genetic evidence, associations with cardio-metabolic outcomes have not been supported thus far.

In conclusion, using data from the largest existing genetic consortia, we demonstrate that genetically lowered 25OHD levels are associated with an increase in the risk of MS in people of European descent. These findings provide rationale for further investigating the potential therapeutic benefits of vitamin D supplementation in preventing the onset and progression of MS.

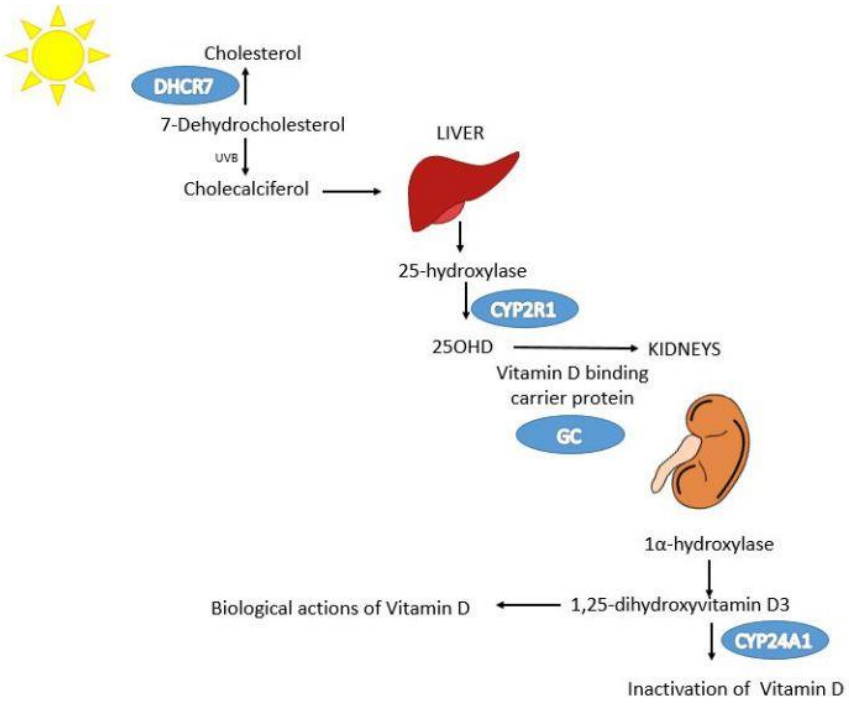
FIGURES

Fig. 2: Schematic representation of Mendelian Randomization analysis



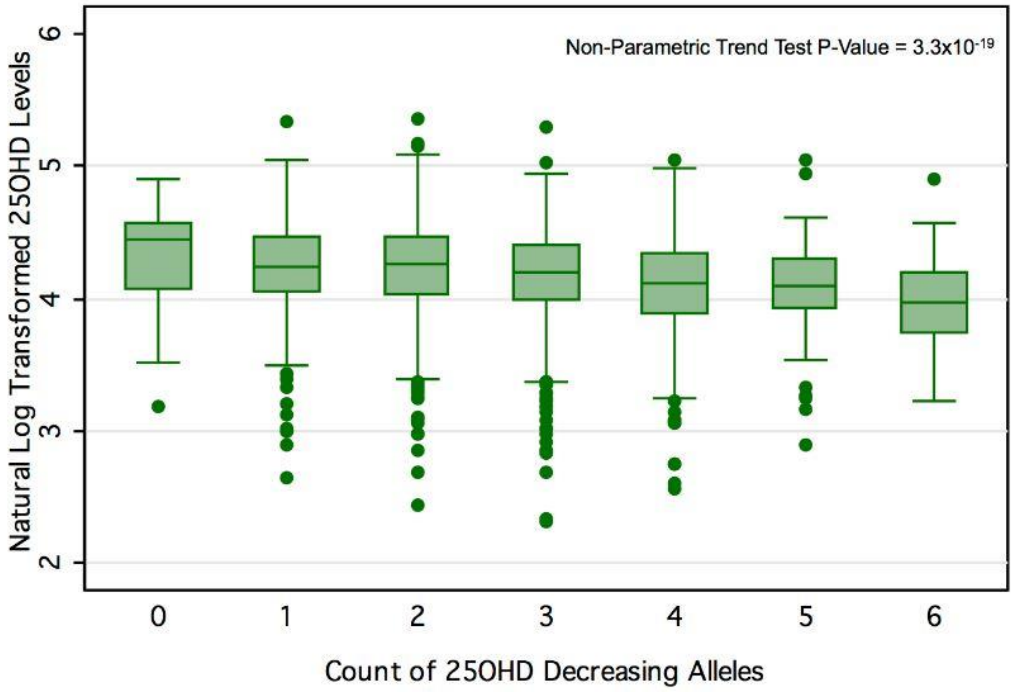
The red box describes SNPs which were genome-wide significant for 25OHD levels in the SUNLIGHT Consortium (n = 33,996). The blue arrow represents the effect of SNPs, on multiply-adjusted, natural log-transformed 25OHD levels using data from the Canadian Multicentre Osteoporosis Study (CaMos, n = 2,347). The green arrow represents the causal association of decreased 25OHD levels on the risk of MS using data from the largest genetic association study to date for MS (the Immunochip study of the International Multiple Sclerosis Genetics Consortium [IMGSC, up to 14,498 cases and 24,091 healthy controls]).

Fig. 3: Vitamin D pathway



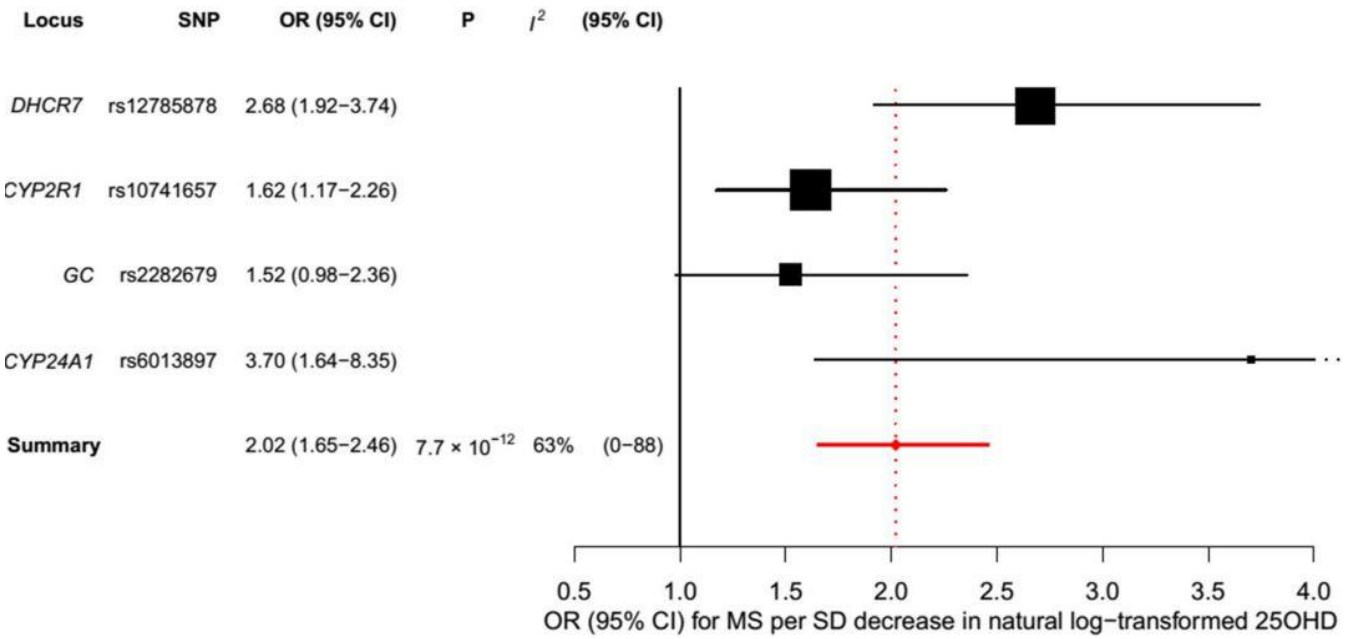
In blue are the genes containing, or in proximity to, SNPs which were genome-wide significant for 25OHD levels in the SUNLIGHT Consortium (n = 33,996). The P-values for association with 25OHD levels were 10^{-109} for *GC,* 10^{-27} for *DHCR7*, 10^{-20} for *CYP2R1* and 10^{-10} for *CYP24A1*. Note that each gene plays an independent role in modulating levels of 25OHD.

Fig. 4: 25OHD Levels by Count of 25OHD Decreasing Alleles in the CaMos cohort



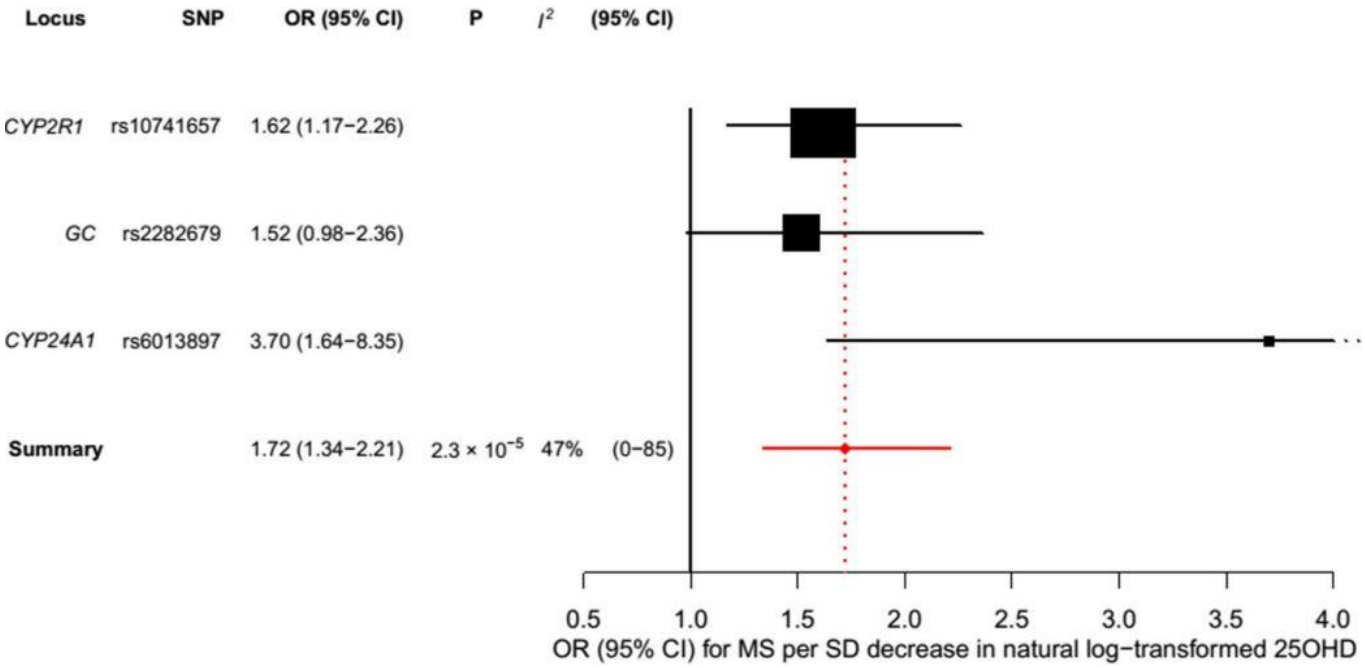
Here we show the box-plot of natural log transformed 25OHD by count of 25OHD decreasing alleles in the CaMos population. A count of zero represents individuals with no 25OHD decreasing alleles (or homozygous at each loci for the 25OHD increasing allele) and a count of six represents an individual with six 25OHD decreasing alleles. No individuals with a count of 7 decreasing alleles or higher was observed in this cohort. The center line and error bars represent the mean levels of natural log transformed 25OHD and their 95% CI for each respective allele count. Note a negative trend between allele count and mean natural log transformed 25OHD.

Fig. 5: Mendelian Randomization Estimate of the Association of 25OHD Levels with Risk of MS



Estimates obtained from using a fixed-effects model

Fig. 6: Mendelian Randomization Estimate of the Association of 25OHD Levels with Risk of MS Excluding the *DHCR7* Locus



Estimates obtained using a fixed-effects model

ACKNOWELDGEMENTS

We wish to kindly thank the SUNLIGHT Consortium, the International Multiple Sclerosis Genetics Consortium and the IMSGC/Wellcome Trust Case Control Consortium 2 study for access to their data.

AUTHOR CONTRIBUTIONS

JBR conceived the experiment. JBR & LEM conducted the analyses. JBR & LEM wrote the first draft of the manuscript. SR created all figures. LEM, JBR, SR OSA, VF, GDS, AL, CMTG & GT wrote the paper. ICMJE criteria for authorship read and met: LEM, JBR, SR, OSA, VF, GDS, AL, CMTG & GT. Agree with manuscript and results: LEM, JBR, SR, OSA, VF, GDS, AL, CMTG & GT.

FUNDING

This work was supported by the Canadian Institute of Health Research, the Canadian Foundation for Innovation, The Fonds de la Recherche en Santé Québec, the Lady Davis Institute and the Jewish General Hospital. AL is supported by the Canadian Diabetes Association Postdoctoral Fellowship

URL: http://www.cihr-irsc.gc.ca/e/193.html

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

COMPETING INTERESTS

GDS is a member of the Editorial Board of PLOS Medicine.

The authors declare no other competing interests exist.

Chapter 4 – Conclusion

In conclusion, our MR study provides strong evidence supporting the effect of vitamin D on the risk of MS, helping to clarify its causal role in MS etiology. Our results have important public health implications, of particular concern to Northern latitude populations that experience both a decreased exposure to sunlight and an increased burden of MS.

The application of MR to this paradigm has important advantages. First, MR greatly diminishes the possibility of confounding, since genotype is unrelated to other self-selected healthy behaviours that influence vitamin D and may independently affect risk of MS. In addition, by employing an MR design, we were able to test the effect of vitamin D in up to 14,498 MS cases, far exceeding the number previously studied through observational or RCT analyses (**Table 1**)**.** We also measured 25OHD, which is a more reliable indicator of vitamin D status than self-reported vitamin D intake as was used in the NHS. Since our results represent lifetime risk of MS due to genetically decreased 25OHD, and thus unlikely to be investigated by RCTs, our results may continue to represent best evidence for vitamin D's causal role in MS incidence. Trials investigating the effect of vitamin D supplementation on MS progression are more likely, but again these studies answer a different research question.

While MR offers numerous advantages, it also has some important limitations. First, pleiotropy, which can introduce bias similar to confounding in observational studies, cannot be fully assessed since the function of many SNPs are unknown. Fortunately, the four SNPs used in our analysis all map to genes involved with known vitamin D mechanisms, therefore it is unlikely that they operate entirely distinct of the vitamin D

pathway. In addition, our analysis cannot address whether compensatory mechanisms attenuate the effect of genetically lowered vitamin D. However, the presence of such mechanisms would likely cause our analysis to under-estimate the true effect of vitamin D on MS risk. By employing the two-sample MR method, we were limited to modeling a linear relationship between vitamin D and MS, which may not be the best fitting model. Rather, there may be a threshold vitamin D level where additional risk incurred due to this genetic susceptibility is abated. This threshold cannot be ascertained through MR studies, only RCTs can address this question. Lastly, we cannot conclude whether vitamin D influences disease progression, therefore our findings are of greatest importance to individuals at high-risk of MS, than those already diagnosed with MS.

More broadly, the implementation of MR is not always feasible since many environmental risk factors do not have sufficiently strong genetic components as required by MR investigation. In addition, an important disadvantage of the two-sample MR is that through the use of summary-statistics, researchers are reliant upon previous GWAS investigators to adjust for proper covariates. This carries consequences as a recent analysis has shown that improper adjustment for heritable traits can introduce bias into genetic effect sizes(99) which could subsequently impact MR analyses.

Only large, long-term RCTs can fully address the efficacy of vitamin D supplementation for the prevention of MS. Similarly, the identification of the clinically important vitamin D level and correct supplement dosage for individuals at high-risk of MS, can only be ascertained through trials. Our multivariate regression model in CaMos estimated that supplementation with ≥ 800 IU increased 25OHD levels by 24 nmol/L relative to supplementation with < 400 IU. This change roughly corresponds to an SD on the

natural-log scale as assessed by our MR study (**Table 6**). However, again this is speculative and only an RCT can assess the proper dose. The IOM also highlights this lack of trial evidence in their 2011 review, calling for more RCTs to identify the correct supplement cut-off points for vitamin D sufficiency.(53) Currently, IOM considers a 25OHD level of 50 nmol/L as sufficient to satisfy the needs of 97.5% of the population.(53) This corresponds to a daily vitamin D intake of 600 IU.(53) The IOM was dismissive of thresholds above these dosages due to the lack of causal evidence for the role of vitamin D outside of skeletal health. However, given our new evidence demonstrating vitamin D's role in MS etiology, this may need to be re-examined with specific recommendations for individuals at high-risk for the disease. Although further research is required, our study provides strong evidence to promote vitamin D awareness among individuals at high-risk of MS.

In conclusion, our MR analysis provides evidence demonstrating vitamin D as a susceptibility factor for MS. While certain questions remain with regards to efficacy and dosage, our study provides evidence to increase vitamin D awareness among individuals at high-risk for MS. Thus MR is a useful tool that can help guide prevention efforts and inform clinical practice.

References

1. Munger KL, Zhang SM, O'Reilly E, Hernan MA, Olek MJ, Willett WC, et al. Vitamin D intake and incidence of multiple sclerosis. Neurology. 2004;62(1):60–5.

2. James E, Dobson R, Kuhle J, Baker D, Giovannoni G, Ramagopalan S V. The effect of vitamin D-related interventions on multiple sclerosis relapses: a meta-analysis. Mult Scler. England; 2013 Oct;19(12):1571–9.

3. Bowling A. Vitamin D and MS: Implications for Clinicians [Internet]. Available from: http://www.nationalmssociety.org/NationalMSSociety/media/MSNationalFiles/Brochures/Clinical-Bulletin-Vitamin-D-and-MS_-Implications-for-Clinical-Practice.pdf

4. Looker A, Pfeiffer C. Serum 25-hydroxyvitamin D status of the US population: 1988–1994 compared with 2000–2004. Am J …. 2008;25(5):1988–94.

5. Yetley EA. Assessing the vitamin D status of the US population. Am J Clin Nutr. United States; 2008 Aug;88(2):558S – 564S.

6. Hartung DM, Bourdette DN, Ahmed SM, Whitham RH. The cost of multiple sclerosis drugs in the US and the pharmaceutical industry: Too big to fail? Neurology. 2015 May 26;84(21):2185–92.

7. Smith GD, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32(1):1–22.

8. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658–65.

9. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet. 2012;380(9841):572–80.

10. Timpson NJ, Lawlor D a., Harbord RM, Gaunt TR, Day INM, Palmer LJ, et al. C-reactive protein and its role in metabolic syndrome: mendelian randomisation study. Lancet. 2005;366(9501):1954–9.

11. Holmes M V, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP, et al. Mendelian randomization of blood lipids for coronary heart disease. Eur Heart J. 2014;1–27.

12. Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. Nat Genet. 2013;45(11):1345–52.

13. Mokry LE, Ahmad O, Forgetta V, Thanassoulis G, Richards JB. Mendelian randomisation applied to drug development in cardiovascular disease: a review. J Med Genet. 2014;jmedgenet – 2014–102438.

14. MS International Federation [Internet]. Available from: http://www.msif.org

15. Ascherio A, Munger KL, White R, Köchert K, Simon KC, Polman CH, et al. Vitamin D as an early predictor of multiple sclerosis activity and progression. JAMA Neurol. 2014;71(3):306–14.

16. Mowry EM, Waubant E, McCulloch CE, Okuda DT, Evangelista A a., Lincoln RR, et al. Vitamin D status predicts new brain magnetic resonance imaging activity in multiple sclerosis. Ann Neurol. 2012;72(2):234–40.

17. Munger KL, Levin LI, Hollis BW, Howard NS, Ascherio A. Serum 25-Hydroxyvitamin D Levels and Risk of Multiple Sclerosis. 2006;296(23):2832–8.

18. Hernán MA, Olek MJ, Ascherio A. Geographic variation of MS incidence in two prospective studies of US women. Neurology. 1999;53(8):1711–8.

19. Kurtzke JF. Epidemiologic contributions to multiple sclerosis: an overview. Neurology. 1980 Jul;30(7 Pt 2):61–79.

20. Simpson S, Blizzard L, Otahal P, Van der Mei I, Taylor B. Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. J Neurol Neurosurg Psychiatry. 2011;82(10):1132–41.

21. Kinoshita M, Obata K, Tanaka M. Latitude has more significant impact on prevalence of multiple sclerosis than ultraviolet level or sunshine duration in Japanese population. Neurol Sci. 2015 Jul;36(7):1147–51.

22. Risco J, Maldonado H, Luna L, Osada J, Ruiz P, Juarez A, et al. Latitudinal prevalence gradient of multiple sclerosis in Latin America. Mult Scler. 2011 Sep;17(9):1055–9.

23. Simpson Jr. S, Blizzard L, Otahal P, Van der Mei I, Taylor B. Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. J Neurol Neurosurg Psychiatry. 2011/04/12 ed. 2011;82(10):1132–41.

24. Holick MF. Sunlight and vitamin D for bone health and prevention of autoimmune diseases, cancers, and cardiovascular disease. Am J Clin Nutr. 2004;80(6 Suppl):1678S – 88S.

25. SWANK RL, LERSTAD O, STRØM A, BACKER J. Multiple sclerosis in rural Norway its geographic and occupational incidence in relation to nutrition. N Engl J Med. 1952 May 8;246(19):722–8.

26. Kampman MT, Wilsgaard T, Mellgren SI. Outdoor activities and diet in childhood and adolescence relate to MS risk above the Arctic Circle. J Neurol. 2007 Mar 21;254(4):471–7.

27. Piantadosi S, Byar DP, Green SB. The ecological fallacy. Am J Epidemiol. 1988 May;127(5):893–904.

28. Kingwell E, Marriott JJ, Jetté N, Pringsheim T, Makhani N, Morrow SA, et al. Incidence and prevalence of multiple sclerosis in Europe: a systematic review. BMC Neurol. 2013 Jan;13:128.

29. Poppe AY, Wolfson C, Zhu B. Prevalence of multiple sclerosis in Canada: a systematic review. Can J Neurol Sci. 2008 Nov;35(5):593–601.

30. Munger KL, Chitnis T, Frazier AL, Giovannucci E, Spiegelman D, Ascherio A. Dietary intake of vitamin D during adolescence and risk of multiple sclerosis. J Neurol. 2011 Mar;258(3):479–85.

31. Kragt J, van Amerongen B, Killestein J, Dijkstra C, Uitdehaag B, Polman C, et al. Higher levels of 25-hydroxyvitamin D are associated with a lower incidence of multiple sclerosis only in women. Mult Scler. 2009 Jan;15(1):9–15.

32. Little RJA, Rosenbaum PR, States. U, Agency. EP, Programs. O of R,

Mathematics. D of S and A. Confounding and selection bias in case control studies. Washington, D.C.: Division of Statistics and Applied Mathematics, Office of Radiation Programs, U.S. Environmental Protection Agency; 1981.

33. Ascherio A, Munger KL, White R, Köchert K, Simon KC, Polman CH, et al. Vitamin D as an Early Predictor of Multiple Sclerosis Activity and Progression. JAMA Neurol. 2014 Mar 1;71(3):306.

34. Mowry EM, Waubant E, McCulloch CE, Okuda DT, Evangelista AA, Lincoln RR, et al. Vitamin D status predicts new brain magnetic resonance imaging activity in multiple sclerosis. Ann Neurol. 2012;72(2):234–40.

35. Runia Tf HWC de RYBBDHRQ, Runia TF, Hop WCJ, De Rijke YB, Buljevac D, Hintzen RQ. Lower serum vitamin D levels are associated with a higher relapse risk in multiple sclerosis. Neurology. 2012;79(3):261.

36. Simpson S, Taylor B, Blizzard L, Ponsonby A-L, Pittas F, Tremlett H, et al. Higher 25-hydroxyvitamin D is associated with lower relapse risk in multiple sclerosis. Ann Neurol. 2010 Aug;68(2):193–203.

37. van der Mei IAF, Ponsonby A-L, Dwyer T, Blizzard L, Simmons R, Taylor B V, et al. Past exposure to sun, skin phenotype, and risk of multiple sclerosis: case-control study. BMJ. 2003 Aug 9;327(7410):316.

38. Langer-Gould A, Brara SM, Beaber BE, Zhang JL. Incidence of multiple sclerosis in multiple racial and ethnic groups. Neurology. 2013 May 6;80(19):1734–9.

39. Willer CJ, Dyment DA, Sadovnick AD, Rothwell PM, Murray TJ, Ebers GC. Timing of birth and risk of multiple sclerosis: population based study. BMJ. 2005 Jan 15;330(7483):120.

40. Bäärnhielm M, Olsson T, Alfredsson L. Fatty fish intake is associated with decreased occurrence of multiple sclerosis. Mult Scler. 2014 May;20(6):726–32.

41. Shrank WH, Patrick AR, Brookhart MA. Healthy user and related biases in observational studies of preventive interventions: a primer for physicians. J Gen Intern Med. 2011 May;26(5):546–50.

42. Vimaleswaran KS, Berry DJ, Lu C, Tikkanen E, Pilz S, Hiraki LT, et al. Causal

Relationship between Obesity and Vitamin D Status: Bi-Directional Mendelian Randomization Analysis of Multiple Cohorts. PLoS Med. 2013;10(2):e1001383.

43. Munger KL, Bentzen J, Laursen B, Stenager E, Koch-Henriksen N, Sørensen TI a, et al. Childhood body mass index and multiple sclerosis risk: a long-term cohort study. Mult Scler. 2013;19(10):1323–9.

44. Munger KL, Chitnis T, Ascherio A. Body size and risk of MS in two cohorts of US women. Neurology. 2009;73(19):1543–50.

45. Burton JM, Kimball S, Vieth R, Bar-Or A, Dosch H-M, Cheung R, et al. A phase I/II dose-escalation trial of vitamin D3 and calcium in multiple sclerosis. Neurology. United States; 2010 Jun;74(23):1852–9.

46. Stein MS, Liu Y, Gray OM, Baker JE, Kolbe SC, Ditchfield MR, et al. A randomized trial of high-dose vitamin D2 in relapsing-remitting multiple sclerosis. Neurology. 2011 Oct 25;77(17):1611–8.

47. Soilu-Hänninen M, Aivo J, Lindström B-M, Elovaara I, Sumelahti M-L, Färkkilä M, et al. A randomised, double blind, placebo controlled trial with vitamin D3 as an add on treatment to interferon β-1b in patients with multiple sclerosis. J Neurol Neurosurg Psychiatry. 2012 May;83(5):565–71.

48. Shaygannejad V, Janghorbani M, Ashtari F, Dehghan H. Effects of Adjunct Low-Dose Vitamin D on Relapsing-Remitting Multiple Sclerosis Progression: Preliminary Findings of a Randomized Placebo-Controlled Trial. Multiple Sclerosis International. 2012. p. 1–7.

49. Kampman MT, Steffensen LH, Mellgren SI, Jorgensen L. Effect of vitamin D3 supplementation on relapses, disease progression, and measures of function in persons with multiple sclerosis: exploratory outcomes from a double-blind randomised controlled trial. Mult Scler J. 2012;18(1):1144–51.

50. Pozuelo-Moyano B, Benito-León J, Mitchell AJ, Hernández-Gallego J. A systematic review of randomized, double-blind, placebo-controlled trials examining the clinical efficacy of vitamin D in multiple sclerosis. Neuroepidemiology. 2013 Jan;40(3):147–53.

51. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of

studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009 Oct;62(10):e1–34.

52. Dörr J, Ohlraun S, Skarabis H, Paul F. Efficacy of vitamin D supplementation in multiple sclerosis (EVIDIMS Trial): study protocol for a randomized controlled trial. Trials. 2012 Jan;13:15.

53. Ross AC. Dietary Reference Intakes for Calcium and Vitamin D - Institute of Medicine [Internet]. Available from: http://www.iom.edu/Reports/2010/Dietary-Reference-Intakes-for-Calcium-and-Vitamin-D.aspx

54. Wang TJ, Zhang F, Richards JB, Kestenbaum B, Van Meurs JB, Berry D, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. Lancet. 2010;376(9736):180–8.

55. Berry DJ, Vimaleswaran KS, Whittaker JC, Hingorani AD, Hyppönen E. Evaluation of genetic markers as instruments for Mendelian randomization studies on vitamin D. PLoS One. 2012;7(5).

56. Sawcer S, Hellenthal G, Pirinen M, Spencer CCA, Patsopoulos NA, Moutsianas L, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011 Aug 11;476(7359):214–9.

57. Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A, Cotsapas C, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45(11):1353–60.

58. Ramagopalan S V., Sadovnick a. D. Epidemiology of Multiple Sclerosis. Neurol Clin. Elsevier Ltd; 2011;29(2):207–17.

59. Hernán M a, Olek MJ, Ascherio a. Geographic variation of MS incidence in two prospective studies of US women. Neurology. 1999;53(8):1711–8.

60. Ebers GC. Environmental factors and multiple sclerosis. Lancet Neurol. England; 2008 Mar;7(3):268–77.

61. Mowry EM, Krupp LB, Milazzo M, Chabas D, Strober JB, Belman AL, et al. Vitamin D status is associated with relapse rate in pediatric-onset multiple sclerosis. Ann Neurol. United States; 2010 May;67(5):618–24.

62. Rosen CJ, Adams JS, Bikle DD, Black DM, Demay MB, Manson JE, et al. The nonskeletal effects of vitamin D: an Endocrine Society scientific statement. Endocr Rev. United States; 2012 Jun;33(3):456–92.

63. Kimball SM, Ursell MR, O'Connor P, Vieth R. Safety of vitamin D3 in adults with multiple sclerosis. Am J Clin Nutr. United States; 2007 Sep;86(3):645–51.

64. Wright S, Yadav V, Bever CJ, Bowen J, Bowling A, Weinstock-Guttman B, et al. Summary of evidence-based guideline: complementary and alternative medicine in multiple sclerosis: report of the Guideline Development Subcommittee of the American Academy of Neurology. Neurology. United States; 2014. p. 1484–6.

65. Looker AC, Pfeiffer CM, Lacher DA, Schleicher RL, Picciano MF, Yetley EA. Serum 25-hydroxyvitamin D status of the US population: 1988-1994 compared with 2000-2004. ajcn.nutrition.org.

66. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133–63.

67. Zacho J, Tybjaerg-Hansen A, Jensen JS, Grande P, Sillesen H, Nordestgaard BG. Genetically elevated C-reactive protein and ischemic vascular disease. N Engl J Med. 2008;359(18):1897–908.

68. Cohen JC, Boerwinkle E, Mosley THJ, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med. United States; 2006 Mar;354(12):1264–72.

69. Faye LL, Sun L, Dimitromanolakis A, Bull SB. A flexible genome-wide bootstrap method that accounts for ranking and threshold-selection bias in GWAS interpretation and replication study design. Stat Med. 2011;30(15):1898–912.

70. Polman CH, Reingold SC, Edan G, Filippi M, Hartung H-P, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria." Ann Neurol. 2005;58(6):840–6.

71. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the

International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001;50(1):121–7.

72. Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol. UNITED STATES; 1983 Mar;13(3):227–31.

73. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2012;491(7422):56–65.

74. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. England; 2014 Sep;23(R1):R89–98.

75. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. Nature Publishing Group; 2011;478(7367):103–9.

76. Dastani Z, Hivert M-FF, Timpson NJ, Perry JRB, Yuan X, Scott R a., et al. Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. PLoS Genet. 2012;8(3):e1002607.

77. Patsopoulos NA, Evangelou E, Ioannidis JPA. Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. Int J Epidemiol. 2008;37(5):1148–57.

78. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557–60.

79. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539–58.

80. Dastani Z, Li R, Richards B. Genetic regulation of vitamin D levels. Calcif Tissue Int. 2013;92(2):106–17.

81. Safadi FF, Thornton P, Magiera H, Hollis BW, Gentile M, Haddad JG, et al.

Osteopathy and resistance to vitamin D toxicity in mice null for vitamin D binding protein. J Clin Invest. 1999;103(2):239–51.

82. Bikle DD, Siiteri PK, Ryzen E, Haddad JG. Serum protein binding of 1,25-dihydroxyvitamin D: a reevaluation by direct measurement of free metabolite levels. J Clin Endocrinol Metab. 1985/11/01 ed. 1985;61(5):969–75.

83. Verboven C, Rabijns A, De Maeyer M, Van Baelen H, Bouillon R, De Ranter C. A structural basis for the unique binding features of the human vitamin D-binding protein. Nat Struct Biol. 2002;9(2):131–6.

84. Shinkyo R, Sakaki T, Kamakura M, Ohta M, Inouye K. Metabolism of vitamin D by human microsomal CYP2R1. Biochem Biophys Res Commun. 2004;324(1):451–7.

85. Kongsbak M, von Essen MR, Levring TB, Schjerling P, Woetmann A, Odum N, et al. Vitamin D-binding protein controls T cell responses to vitamin D. BMC Immunol. 2014/09/19 ed. 2014;15:35.

86. Moon M, Song H, Hong HJ, Nam DW, Cha M-Y, Oh MS, et al. Vitamin D-binding protein interacts with Aβ and suppresses Aβ-mediated pathology. Cell Death Differ. 2013;20(4):630–8.

87. Rinaldi AO, Sanseverino I, Purificato C, Cortese A, Mechelli R, Francisci S, et al. Increased circulating levels of vitamin D binding protein in MS patients. Toxins (Basel). 2015;7(1):129–37.

88. Yang M, Qin Z, Zhu Y, Li Y, Qin Y, Jing Y, et al. Vitamin D-binding protein in cerebrospinal fluid is associated with multiple sclerosis progression. Mol Neurobiol. 2013;47(3):946–56.

89. Browne RW, Weinstock-Guttman B, Zivadinov R, Horakova D, Bodziak M Lou, Tamaño-Blanco M, et al. Serum lipoprotein composition and vitamin D metabolite levels in clinically isolated syndromes: Results from a multi-center study. J Steroid Biochem Mol Biol. 2014;143:424–33.

90. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. United States; 2013 Nov;45(11):1274–83.

91. van Luijn MM, Kreft KL, Jongsma ML, Mes SW, Wierenga-Wolf AF, van Meurs M, et al. Multiple sclerosis-associated CLEC16A controls HLA class II expression via late endosome biogenesis. Brain. 2015;138(Pt 6):1531–47.

92. Ramagopalan S V, Maugeri NJ, Handunnetthi L, Lincoln MR, Orton S-M, Dyment DA, et al. Expression of the multiple sclerosis-associated MHC class II Allele HLA-DRB1*1501 is regulated by vitamin D. PLoS Genet. 2009;5(2):e1000369.

93. O'Connell K, Kelly S, Kinsella K, Jordan S, Kenny O, Murphy D, et al. Dose-related effects of vitamin D on immune responses in patients with clinically isolated syndrome and healthy control participants: study protocol for an exploratory randomized double- blind placebo-controlled trial. Trials. 2013;14:272.

94. Thouvenot E. Efficacy of Cholecalciferol (Vitamin D3) for Delaying the Diagnosis of MS After a Clinically Isolated Syndrome (D-Lay-MS) [Internet]. Available from: https://clinicaltrials.gov/ct2/show/NCT01817166?term=vitamin+d+multiple+sclerosis&rank=13

95. Leong A, Rehman W, Dastani Z, Greenwood C, Timpson N, Langsetmo L, et al. The Causal Effect of Vitamin D Binding Protein (DBP) Levels on Calcemic and Cardiometabolic Diseases: A Mendelian Randomization Study. PLoS Med. 2014;11(10):e1001751.

96. Ye Z, Sharp SJ, Burgess S, Scott RA, Imamura F, Consortium I, et al. Association between circulating 25-hydroxyvitamin D and incident type 2 diabetes: a mendelian randomisation study. Lancet Diabetes Endocrinol. 2015;3(1):35–42.

97. Cooper JD, Smyth DJ, Walker NM, Stevens H, Burren OS, Wallace C, et al. Inherited variation in vitamin D genes is associated with predisposition to autoimmune disease type 1 diabetes. Diabetes. 2011;60(5):1624–31.

98. Vimaleswaran KS, Cavadino A, Berry DJ, investigators LCS, Jorde R, Dieffenbach AK, et al. Association of vitamin D status with arterial blood pressure and hypertension risk: a mendelian randomisation study. Lancet Diabetes Endocrinol. 2014;2(9):719–29.

99. Aschard H, Vilhjálmsson BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. Am J Hum Genet. 2015 Feb 5;96(2):329–39.

Appendices

Additional Regression Models of 25OHD in the CaMos Population

Regression of 25OHD on the Absolute Scale

The examiner of this thesis asked for a comparison of the effect of the SUNLIGHT SNPs to those of season and supplementation on the absolute scale in CaMos. Note that this analysis was not a part of the PLOS Medicine manuscript.

	Univariate*			Multivariate†		
	β	**SE**	**p-value**	**β**	**SE**	**p-value**
SNPs						
rs6013897 (per A allele)	-1.55	0.99	0.12	-0.91	0.88	0.30
rs2282679 (per C allele)	-3.66	0.84	1.23E-05	-3.51	0.74	2.53E-06
rs10741657 (per C allele)	-3.61	0.80	5.97E-06	-3.61	0.71	4.17E-07
rs12785878 (per G allele)	-2.85	0.87	0.0010	-3.48	0.78	8.29E-06
Supplementation						
< 400 IU (ref)	-	-	-	-	-	-
400 - 800 IU	10.56	1.07	< 2E-16	12.00	1.13	< 2E-16
≥ 800 IU	22.42	1.29	< 2E-16	24.45	1.48	< 2E-16
Age						
Age (years)	-0.06	0.04	0.20	-0.10	0.05	0.029
Age^2	0.01	0.00	6.08E-05	0.01	0.00	0.0014
Sex						
Male (ref)	-	-	-	-	-	-
Female	-0.78	1.06	0.46	-4.78	1.08	1.03E-05
BMI						
BMI (kg/m2)	-0.96	0.09	<2E-16	-0.82	0.09	< 2E-16
Season						
January-March (ref)	-	-	-	-	-	-
-April-June	4.67	1.33	0.00046	5.37	1.35	7.33E-05
July-September	12.18	1.43	< 2E-16	11.50	1.47	7.25E-15
October-December	6.61	1.44	4.56E-06	6.21	1.46	2.11E-05

* Univariate is a regression model including only one independent variable upon 25OHD. This was done to assess for possible confounding

† Multivariate is a fully adjusted regression model including all four SUNLIGHT SNPs, age, age^2 sex, BMI, season and supplementation as covariates

Using a Categorical Variable for Month versus Season

The examiner of this thesis also noted that vitamin D levels vary greatly throughout the year and questioned our approach collapsing of months into 4 seasons (January-March, April-June, July-September, October-December). To explore this suggestion we performed a regression of month on 25OHD in CaMos

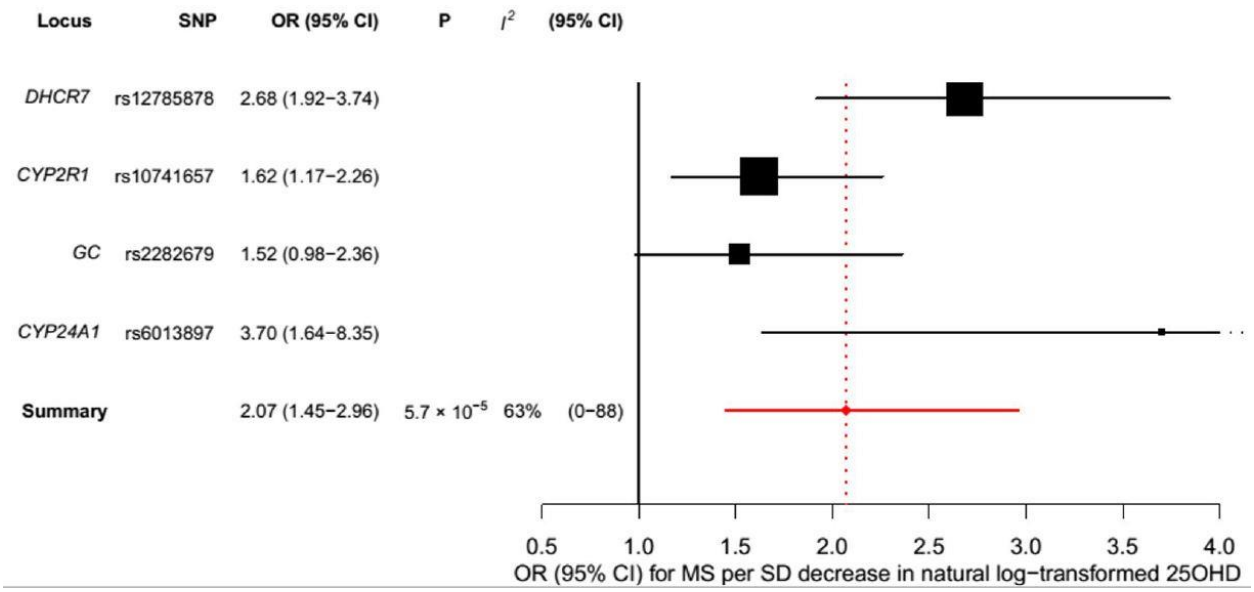
Month	β	SE	p-value
January (ref)	-	-	-
February	-4.00	2.57	0.12
March	-8.87	2.48	0.00034
April	-2.36	2.32	0.31
May	0.61	2.26	0.79
June	2.37	2.31	0.30
July	5.88	2.45	0.016
August	8.49	2.48	0.00062
September	8.99	2.47	0.00028
October	5.88	2.34	0.012
November	-0.72	2.49	0.77
December	-0.43	2.71	0.87

Upon further assessment of month as an independent variable, we believe that this validates our initial approach. While vitamin D measurements are likely to vary on a monthly basis, as noted by the standard error, in the CaMos cohort there are too few measurements in any given month to estimate its effect accurately. This leads to results that lack a clear interpretation as 25OHD measurement taken during some months, such as August and September, strongly increase 25OHD levels as compared to January, whereas June and May do not. By using season, this allows us both to increase the number of observations falling within each category and to capture the accumulative effect of season (and sun exposure) on 25OHD levels.

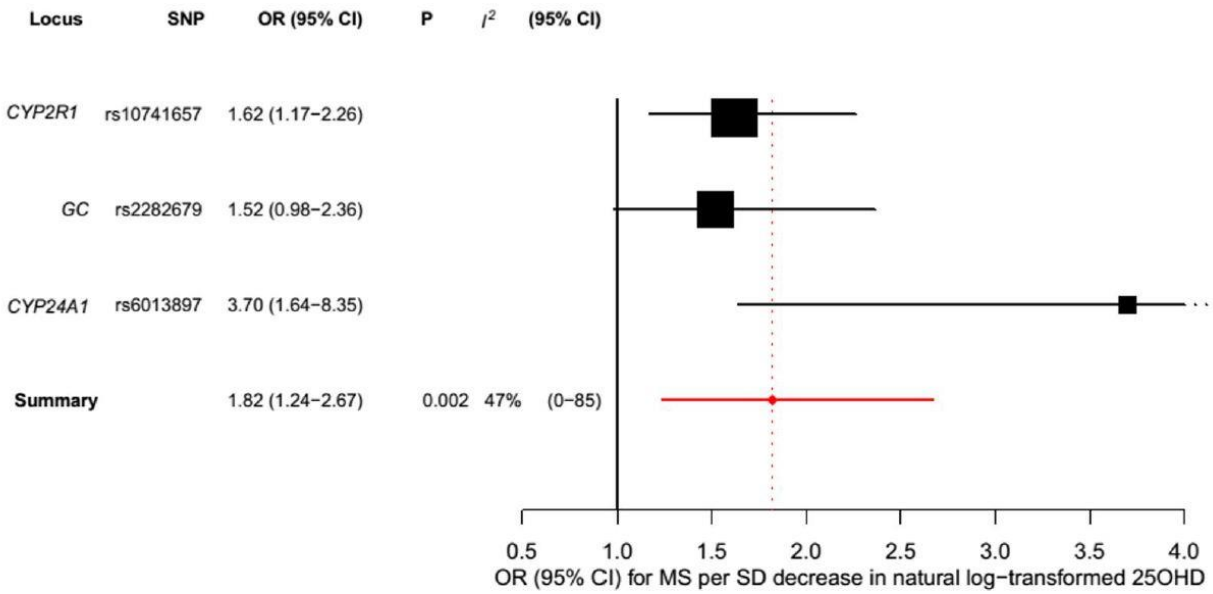
Supplementary Information for PLOS Medicine Manuscript

Supplementary Figures

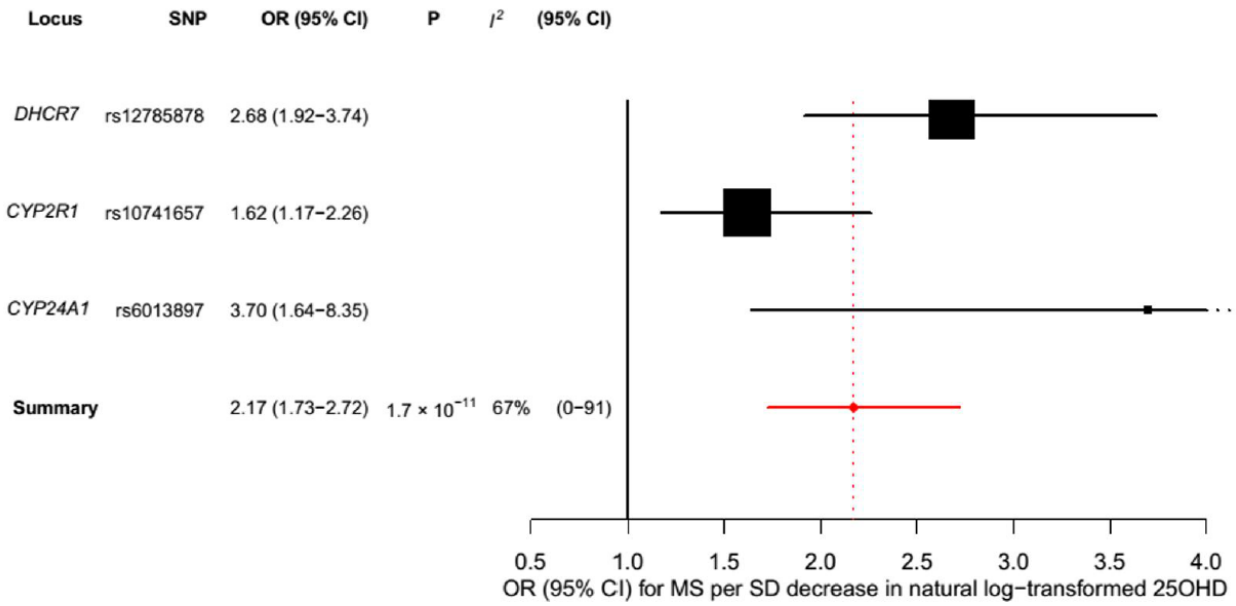
S1 Fig. MR estimate of the association of 25OHD levels with risk of MS using a random effects model



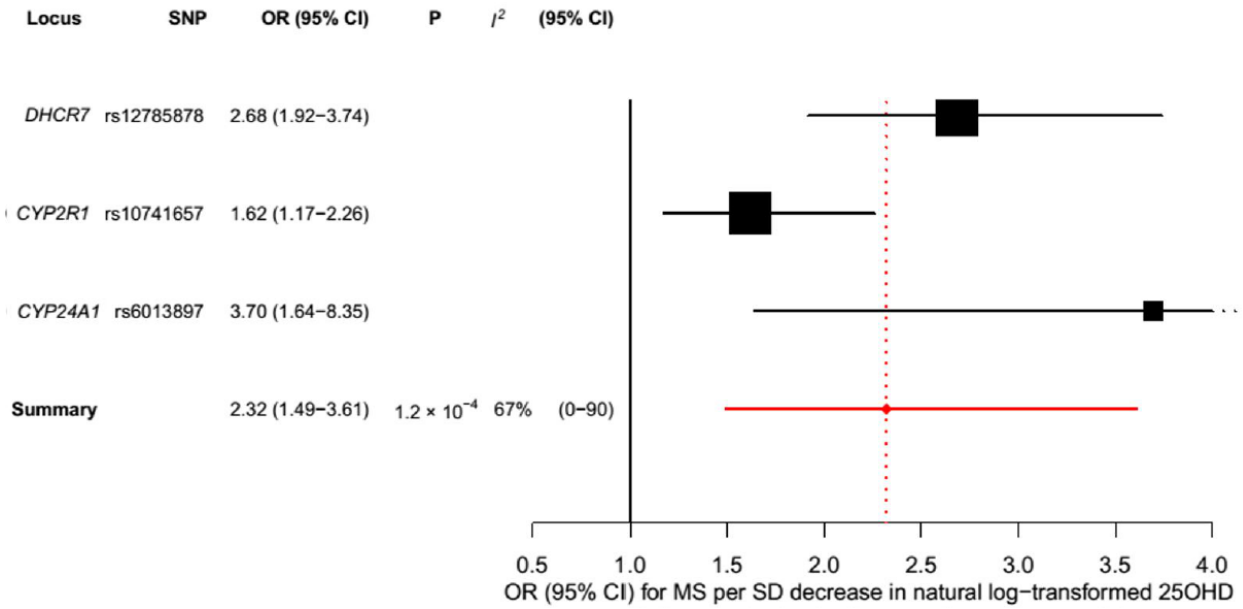
S2 Fig. MR estimate of the association of 25OHD levels with risk of MS excluding the DHCR7 locus using a random effects model



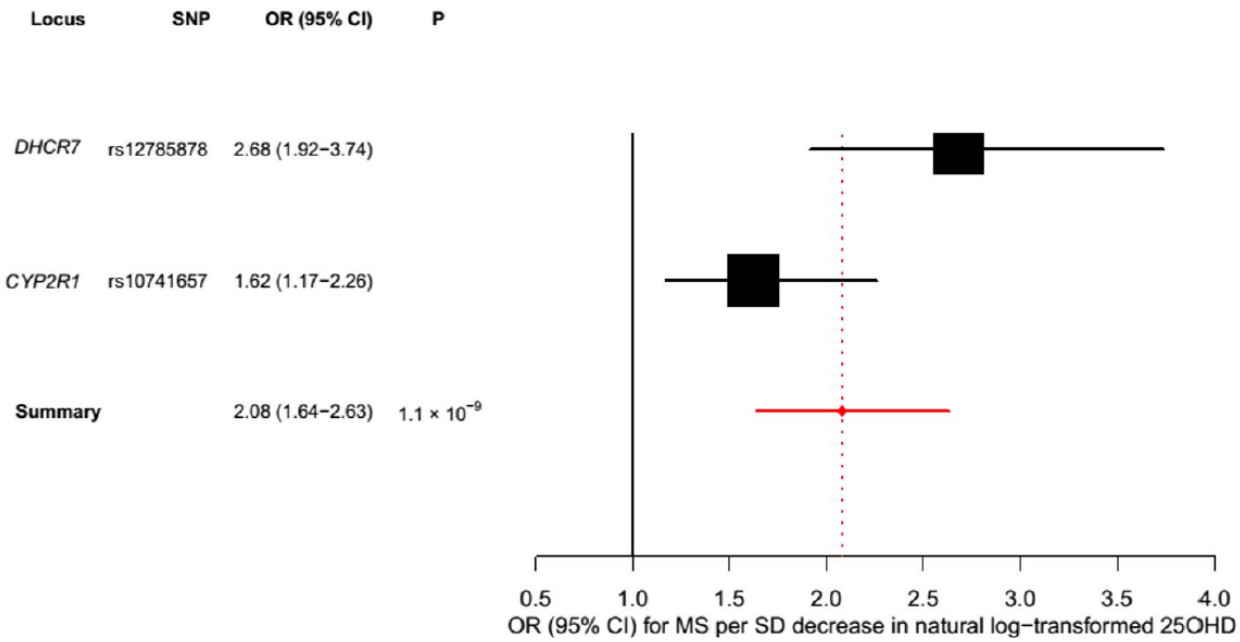
S3 Fig. MR estimate of the association of 25OHD levels with risk of MS excluding the *GC* locus using a fixed effects model



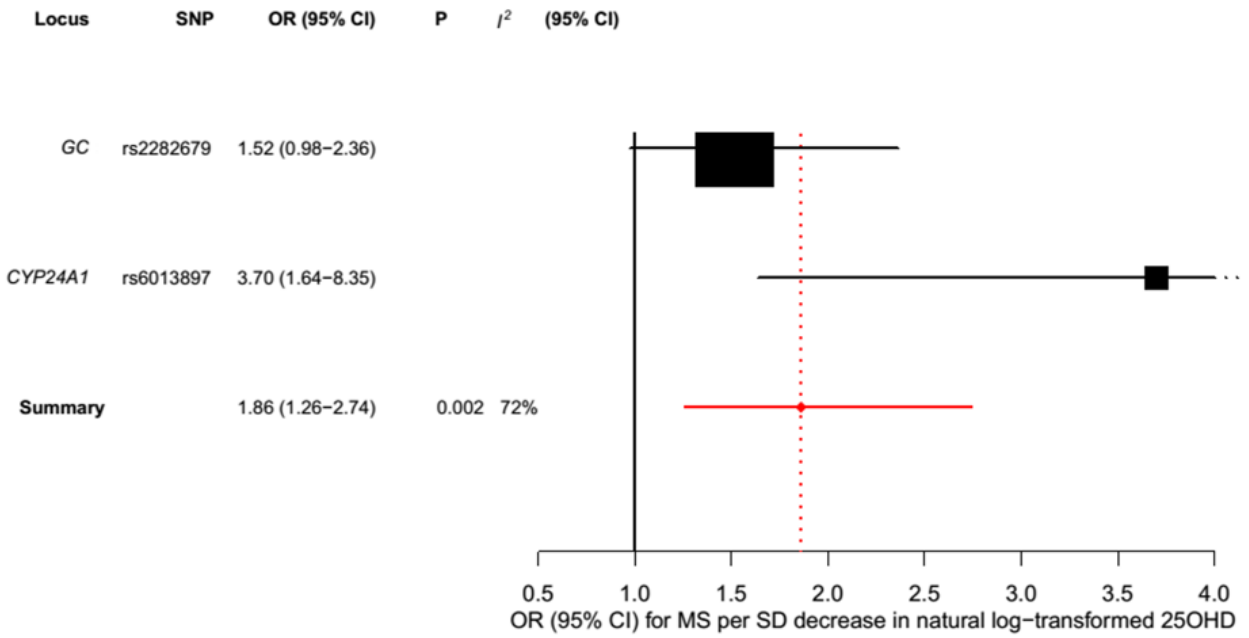
S4 Fig. MR estimate of the association of 25OHD levels with risk of MS excluding the *GC* locus using a random effects model



S5 Fig. MR estimate of the association of 25OHD levels with risk of MS for SNPs involved in 25OHD synthesis using a fixed effects model



S6 Fig. MR estimate of the association of 25OHD levels with risk of MS for SNPs involved in 25OHD metabolism using a fixed effects model



S1 Table

Summary of previous work assessing the association between 25OHD SNPs and relevant biomarker pathways (taken from previous work by Berry at al.) (55)

Biomarkers	P-Values			
	GC	*DHCR7*	*CYP24A1*	*CYP2R1*
vWF	0.43	0.30	0.97	0.10
tPA	0.51	0.17	0.27	0.34
D-dimer	0.80	0.66	0.63	0.65
Fibrinogen	0.66	0.90	0.61	0.40
CRP	0.04	0.44	0.08	0.99
Triglycerides	0.81	0.38	0.54	0.47
LDL	0.50	0.07	0.62	0.28
HDL	0.94	0.62	0.52	0.95
Cholesterol	0.73	0.10	0.70	0.39
FEV	0.56	0.95	0.76	0.20
Diastolic BP	0.17	0.34	0.065	0.12
Systolic BP	0.26	0.30	0.03	0.89
IgE	0.59	0.75	0.60	0.03
IGF-1	0.06	0.90	0.84	0.70
HbA1c	0.56	0.94	0.72	0.90

S1 Methods

PubMed Search

The following terms were searched on the PubMed database to investigate possible pleiotropic mechanisms of our chosen SNPs corresponding to gene name, gene mutations, encoded protein, encoded protein with MS and encoded protein with autoimmunity.

For rs2282679: "*GC", "GC gene", "GC gene mutations", "vitamin D binding protein", "vitamin D binding protein multiple sclerosis", "vitamin D binding protein autoimmunity",*

The search term GC uncovered 69152 results, most of which were not relevant to genetics, therefore the search term "GC gene" was used instead to refine search results.

For rs12785878: "*DHCR7", "DHCR7 mutations", "7-dehydrocholesterol reductase", "7-dehydrocholesterol reductase multiple sclerosis", 7-dehydrocholesterol reductase autoimmunity"*

For rs6013897: "*CYP24A1", "CYP24A1 mutations", "1,25-dihydroxyvitamin D3 24-hydroxylase", "1,25-dihydroxyvitamin D3 24-hydroxylase multiple sclerosis" "1,25-dihydroxyvitamin D3 24-hydroxylase autoimmunity"*

For rs10741657: "*CYP2R1", "CYP2R1 mutations", "vitamin-D hydroxylase", "vitamin-D hydroxylase multiple sclerosis", "vitamin-D hydroxylase autoimmunity".*

Abstracts were selected for further review if they made reference to the search term and a pathway distinct from vitamin D or vitamin D insufficiency/ deficiency on MS or

autoimmunity. Only studies in mammals were considered. Vitamin-D related genes have often been associated with other clinical phenotypes such as colorectal cancer and inflammatory airway conditions however these were considered unrelated to MS and autoimmunity. Findings are reported in the Results section.