UC Berkeley
UC Berkeley Previously Published Works

Title

Evidence supports a causal association between allele-specific vitamin D receptor binding and multiple sclerosis among Europeans.

Permalink

Journal

Authors

Adams, Cameron
Manouchehrinia, Ali
Quach, Hong
et al.

Publication Date

DOI

Peer reviewed

Evidence supports a causal association between allele-specific vitamin D receptor binding and multiple sclerosis among Europeans

Cameron Adams[a,1], **Ali Manouchehrinia**[b,g,1], **Hong L. Quach**[a], **Diana L. Quach**[a], **Tomas Olsson**[b,f,g], **Ingrid Kockum**[b,f,g], **Catherine Schaefer**[e], **Chris P. Ponting**[d], **Lars Alfredsson**[b,c,h], **Lisa F. Barcellos**[a,e,2]

[a]Genetic Epidemiology and Genomics Laboratory, School of Public Health, University of California, Berkeley, CA, USA

[b]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

[c]Centre for Occupational and Environmental Medicine, Region Stockholm, Stockholm, Sweden

[d]MRC Human Genetics Unit, The Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, UK

[e]Kaiser Permanente Division of Research, Kaiser Permanente Northern California, Oakland, California, USA

[f]Academic Specialist Center, Stockholm, Sweden

[g]Centrum for molecular medicine, Karolinska University Hospital, Stockholm, Sweden

[h]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

Abstract

Although evidence exists for a causal association between 25-hydroxyvitamin D (25(OH)D) serum levels and multiple sclerosis (MS), the role of variation in vitamin D receptor (VDR) binding in MS is unknown. Here, we leveraged previously identified variants associated with allele imbalance in VDR binding (VDR-binding variant; VDR-BV) in ChIP-exo data from calcitriol-stimulated lymphoblastoid cell lines and 25(OH)D serum levels from genome-wide association studies to construct genetic instrumental variables (GIVs). GIVs are composed of one or more genetic variants that serve as proxies for exposures of interest. Here, GIVs for both VDR-BVs and 25(OH)D were used in a two-sample Mendelian Randomization study to investigate the relationship between VDR binding at a locus, 25(OH)D serum levels, and MS risk. Data for 13,598 MS cases and 38,887 controls of European ancestry from Kaiser Permanente Northern California, Swedish MS studies, and the UK Biobank were included. We estimated the association between each VDR-BV GIV and MS. Significant interaction between a VDR-BV

GIV and a GIV for serum 25OH(D) was evidence for a causal association between VDR-BVs and MS unbiased by pleiotropy. We observed evidence for associations between two VDR-BVs (rs2881514, rs2531804) and MS after correction for multiple tests. There was evidence of interaction between rs2881514 and a 25(OH)D GIV, providing evidence of a causal association between rs2881514 and MS. This study is the first to demonstrate evidence that variation in VDR binding at a locus contributes to MS risk. Our results are relevant to other autoimmune diseases in which vitamin D plays a role.

Classification

BIOLOGICAL SCIENCES; Genetics

Introduction

Multiple sclerosis (MS) is more prevalent among people residing in latitudes farther from the equator (1). This has led to the hypothesis that low vitamin D levels are associated with increased MS susceptibility (2). In observational studies, low levels of serum 25-hydroxyvitamin D (25(OH)D) (3) and lack of sunlight exposure (4, 5) are associated with an increased risk of MS, and vitamin D supplements and diets rich in vitamin D (6, 7) are associated with decreased risk of MS. However, the potential for reverse causation and unmeasured confounding factors make it difficult to assess alternative explanations. Studies employing Mendelian Randomization (MR), a method that utilizes genetic variation in genes associated with an exposure or other variables of interest (genetic instrumental variables or GIVs) to estimate a causal association between that exposure and disease(8), strongly support a causal relationship between low serum 25(OH)D levels and increased MS susceptibility (9–11). Many genetic variants used in these MR studies are involved in vitamin D biosynthesis, suggesting that different aspects of the vitamin D pathway, specifically transcription and expression mediated by vitamin D receptor DNA binding, contribute to the relationship between vitamin D and MS (12, 13).

25(OH)D signals through the nuclear vitamin D receptor (VDR), a ligand-regulated transcription factor that mediates all genomic actions of 25(OH)D (Fig. 1A) (14). Upon activation by vitamin D, the VDR forms the RXR/VDR heterodimer that acts as a transcription factor by binding to specific DNA regions across the genome. Most of these regions are vitamin D response elements (VDREs), a particular DNA sequence located in gene promoter regions defined by its high VDR binding affinity; however, not all VDR binding sites contain VDREs (15, 16). Through this pathway, VDR binding modulates the transcription of genes across the genome and 25(OH)D signaling through the VDR can lead to target gene expression (17). Through VDR-mediated gene transcription, vitamin D regulates calcium metabolism, cellular growth, proliferation, apoptosis, and inflammation (18). VDR binds preferentially to one allele over the other at heterozygous sites in Lymphoblastoid Cell-Lines (LCLs) (Fig. 1B), as measured using chromatin immunoprecipitation followed by exonuclease digestion (ChIP-Exo) (19). 1000 Genomes data reveals that VDR-BVs are enriched within genomic regions associated with autoimmune conditions, including MS (19).

Individual VDR-BVs have not been evaluated for association with MS susceptibility in large studies. Here, we leveraged available data on VDR-BVs and results from genome-wide association studies of 25(OH)D serum levels to construct GIVs for both VDR-BV and 25(OH)D serum levels using imputed whole genome SNP data. GIVs can serve as proxies for exposures and other variables of interest (here for VDR-BVs and 25(OH)D) and were used in a two-sample Mendelian Randomization study to investigate the causal relationship between VDR binding at a locus, vitamin D serum levels and MS risk. We also utilized the dependency between VDR binding and bioavailability of 25(OH)D to develop an MR analysis framework to account for horizontal pleiotropy, a source of bias where a GIV (VDR-BV) affects disease risk outside of its effect on the variable of interest. For a true causal association, the bioavailability of 25(OH)D should modulate the effect of genetic variation in VDR binding on MS susceptibility. A significant interaction between the bioavailability of 25(OH)D and VDR binding, captured by the GIVs, would provide evidence of an association between VDR binding at a locus and MS that is not biased by horizontal pleiotropy (Fig.1C). We hypothesized that altered VDR binding modulates transcription and expression of a target gene, thereby increasing or decreasing the risk of MS. We utilized data on individuals from three MS case-control studies and the United Kingdom (UK) Biobank. Identification of VDR-BVs associated with MS could improve understanding of the biological mechanisms through which vitamin D acts to affect MS and further elucidate the molecular and cellular causes of MS.

Results

Characteristics of MS cases and controls

The largest source of MS cases was the Swedish Human OMNI (OMNI; 6,709 MS cases and 5,881 controls), followed by the Swedish GSA (GSA; 3,718 MS cases, 1,180 controls), the UK Biobank study (UKB; 2,087 MS cases, 20,870 matched controls), and Kaiser Permanente Northern California (KPNC; 1,082 cases, 10,956 controls) (Table 1). The individuals in Swedish cohorts were genotyped on two different platforms and analyzed separately. Consistent with the female predominance in MS, all studies were primarily comprised of female participants with KPNC having the highest proportion of females (80.2%) and UKB having the lowest (72.4%). The average birth year was earlier within KPNC and UKB studies (the mid-1950s) than the other studies (1960-1970s). The proportion of MS cases with at least one *HLA-DRB1*15:01* allele was similar across all studies (~49-57%), with UKB cases having the lowest carriage rate. All analyses were performed separately within KPNC, GSA, OMNI, and UKB and combined with meta-analysis.

Linkage Disequilibrium between VDR-BVs and MS GWAS risk variants

Within the European samples from the 1000 Genomes project, linkage disequilibrium (LD) between VDR-BVs and established MS GWAS variants was generally low. Of the 112 VDR-BVs considered in this investigation, only two VDR-BVs had an $r^2 > 0.2$ with a known MS GWAS variant (VDR-BV: rs55792977, MS GWAS: rs13385171, r^2=0.48; VDR-BV: rs13098781, MS GWAS: rs9863496, r^2=0.34), indicating no correlation between MS GWAS risk variants and nearly all VDR-BV variants.

25(OH)D serum levels are casually associated with MS susceptibility

Two GIVs constructed using summary statistics from two recent GWASs on *25(OH)D serum levels by* Jiang et al. and Revez et al. (hereafter, GIV_{25OHD}; see Methods) were used in our analyses (12, 13). The mean $GIV_{25(OH)D}$ for increased 25(OH)D serum levels was lower among MS cases compared to controls in all studies (Table 1). The difference for both the Jiang et al. $GIV_{25(OH)D}$ and Revez et al. $GIV_{25(OH)D}$ means between MS cases and controls was the largest among Sweden participants genotyped using the GSA chip. The difference in the means was smallest across both $GIV_{25(OH)D}$ between UKB MS cases and controls. Consistent with previous research, a decreased 25(OH)D serum level captured by each $GIV_{25(OH)D}$ was associated with an increased risk of MS (Fig. 2A-B). The magnitude of the association was greater for the Jiang et al. $GIV_{25(OH)D}$ (OR: 1.85, 95% CI: 1.30-2.63) compared to the Revez et al. $GIV_{25(OH)D}$ (OR: 1.32, 95% CI: 1.07-1.61), likely owing to the difference in the number of SNPs within each GIV (Methods) and reported effect sizes. Estimates for between study variance were low for the Jiang et al. $GIV_{25(OH)D}$ (I^2=0%, p_Q=0.56) indicating associations were consistent across all studies. While there was some evidence for moderate heterogeneity for the Revez et al. $GIV_{25(OH)D}$ (I^2=55%, p_Q=0.09), the direction of effect was consistent across all studies, providing support for a uniform pattern across the studies, and strong evidence for causal association between low 25(OH)D serum level and MS risk.

VDR-BV *associations with MS*

GIVs for each VDR-BV (hereafter, GIV_{VDR}) were constructed (See Methods). Of the 112 $GIVs_{VDR}$ included in our analyses, seven $GIVs_{VDR}$ were associated with MS at p<0.05 (Table 2). Following correction for multiple testing, two GIV_{VDR} remained significant (p_{FDR}<0.05). Meta-analyses of these two $GIVs_{VDR}$ demonstrated evidence of an association between variation within VDR binding and MS susceptibility across all four studies. GIV_{VDR} rs2881514 was associated with an increased risk of MS (Meta-analysis Odds Ratio (OR): 1.10, 95% CI: 1.05-1.15, p=9.4×10^{-5}; Fig. 2C) and GIV_{VDR} rs2531804 was associated with a decreased risk of MS (Meta-analysis OR: 0.82, 95% CI: 0.73-0.92, p=6.4×10^{-4}). There was evidence of interaction between 13 $GIVs_{VDR}$ and GIV_{25OHD} at p<0.05 (SI Appendix, Table S1). No interactions were significant after correcting for multiple testing; however, GIV_{VDR} rs2881514, the top finding from the independent models, was among these 13 (GIV_{VDR} rs2881514 × Jiang et al. GIV_{25OHD} Meta-analysis OR: 2.17, 95% CI: 1.10-4.29, p=0.025; Fig. 2D). The direction of this association provided evidence of interaction between *increased* VDR binding at GIV_{VDR} rs2881514 and *decreased* bioavailability of 25(OH)D measured with Jiang et al. GIV_{25OHD} conferring *increased* risk of MS.

VDR-BVs with evidence of MS association are eQTLs and sQTLs

Several of the GIV_{VDR} that were independently associated at p<0.05 with MS or showed evidence of interaction with $GIV_{25(OH)D}$ are expression quantitative trait loci (eQTL) or splicing quantitative trait loci (sQTL) for one or more human tissues in GTEx (SI Appendix, Fig. S1) (20). GIV_{VDR} rs2881514 is an eQTL for *RFTN1* in esophageal tissue and an sQTL for *RFTN1* in EBV-transformed lymphocytes and thyroid tissues. *RFTN1* encodes Raftlin, a

B cell-specific major raft protein necessary for lipid raft integrity and B-cell receptor signal transduction (21). Other GIV$_{VDR}$, including rs2531804, rs55811049, rs961320, rs871699, and rs62200158, which showed evidence of association with MS at p<0.05, are eQTLs or sQTLs in brain or skin tissue, among others.

Discussion

This is the first study to observe evidence for associations between variation in VDR binding at a locus and MS susceptibility, providing further evidence of the important role of the vitamin D pathway in MS. Our results support the hypothesis that alterations in VDR binding and subsequent modulations in vitamin D gene transcription and expression contributes to developing MS. We found two GIV$_{VDR}$, rs2881514 and rs2531804, were independently associated with MS susceptibility. There was also evidence of interaction between several GIVs$_{VDR}$ and GIV$_{25(OH)D}$ and MS risk. Although no observed interactions were significant after correction for multiple testing, there was evidence of interaction between GIV$_{VDR}$ rs2881514 and the Jiang et al. GIV$_{25(OH)D}$ at p<0.05. It is important to note that detecting interactions requires substantially more statistical power than main effects. Regardless, evidence of an association between a GIV$_{VDR}$ and MS in independent and interaction models is promising (22). Both GIV$_{25(OH)D}$ for 25(OH)D serum levels were significantly associated with MS; specifically, lower 25(OH)D serum was associated with increased risk of MS in all four studies.

There are causal associations between vitamin D serum levels and MS risk using MR methods (9–11). Our findings using GIVs constructed using summary statistics from two recent GWASs on 25(OH)D serum levels are consistent with previous MR studies and further confirm vitamin D insufficiency as an important risk factor for MS. Measures of heterogeneity indicate little variation between studies for Jiang et al. GIV$_{25(OH)D}$, although there was some evidence of moderate heterogeneity for the Revez et al. GIV$_{25(OH)D}$ (I^2=55%, p_Q=0.09), which might be explained by the larger number of SNPs included in the Revez et al. GIV$_{25(OH)D}$. Associations were modest for both GIV$_{25(OH)D}$ among UKB participants. Unlike the other studies, MS cases from UKB did not have their MS status confirmed by a neurologist. Instead, diagnoses were inferred from electronic health records or self-report and may have resulted in a small amount of misclassification of MS case status in this cohort. Evidence of this may be seen in the lower frequency of carriage of *HLA-DRB1*15:01*, the major risk allele for MS risk, in UKB MS cases (49.7%) compared to cases in the other three studies (>53%).

The top GIV$_{VDR}$ associated with MS from meta-analyses was rs2881514. The A allele of this SNP was associated with increased VDR binding in LCLs. Evidence from results within each study and meta-analyses indicate a harmful effect of increased VDR binding at that locus on MS risk. Further, the interaction observed between GIV$_{VDR}$ rs2881514 and GIV$_{25(OH)D}$ demonstrates evidence for interaction between *increased* VDR binding at rs2881514 and *decreased* 25(OH)D serum levels, yielding a substantial increase in the risk of MS (Meta-analysis OR: 2.17, 95% CI: 1.10-4.29). This association was consistent among UKB, OMNI, and GSA MS cases and controls but was slightly protective in KPNC, though very wide confidence intervals were observed (KPNC OR: 0.79, 95% CI:

0.14-4.47). GIV$_{VDR}$ rs2881514 is located on chromosome 3, 1,533 bases downstream of the transcription start site of *RFTN1. In GTEx*, rs2881514 is an eQTL for *RFTN1* in esophageal tissue and an sQTL for *RFTN1* in EBV-transformed lymphocytes and thyroid tissues, providing evidence for a role of rs2881514 in *RFTN1* expression. *RFTN1* encodes Raftlin, a protein that is critical for producing lipid rafts, which are membrane microdomains that play a crucial role in B cell activation through B cell receptor signaling (21). Raftlin is essential for clathrin-dependent endocytosis of Toll-like Receptor (TLR) 3 ligand in human epithelial cells and myeloid dendritic cells (23), and for LPS-induced TLR4 internalization and Toll-IL-1R domain-containing adaptor molecule-1 (TICAM-1) signaling in human monocyte-derived DCs and macrophages (24). TLR3 and TLR4 signaling pathways have been implicated in MS pathogenesis and symptom modulation (25). In mice, experimental autoimmune encephalomyelitis (EAE), a Th1-/Th17-mediated disease, was more severe among mice with wild-type Raftlin expression compared to Raftlin deficient mice (23). *RFTN1* has not been directly implicated in MS risk; however, there is increased expression of *RFTN1* in chronic active brain lesions from MS cases compared to healthy brain tissue from controls (26). More work is needed to understand the relationship between *RFTN1* and vitamin D, and its role in MS.

Several of the GIV$_{VDR}$ associated with MS had evidence of SNP-associated tissue-specific expression and alternative splicing in GTEx (SI Appendix, Fig. S1). rs25318104, rs9621320, rs871699, rs689384, rs62200158, rs558110449, and rs2286576 show evidence of being an eQTL or sQTL in one or more brain tissues. Many of these SNPs are eQTLs or sQTLs across various tissues. Of the set of VDR-BVs tested as GIV$_{VDR}$ in the current study, those with evidence of association with MS susceptibility were enriched for genes involved in immune-related processes (leukocyte-mediated immunity, immune effector process, myeloid leukocyte activation, and phospholipase activity), vesicles, and exocytosis (SI Appendix, Table S2). Annotations to eQTLs and sQTLs and the enrichment results provide evidence supporting a mediating role of allele-specific VDR binding in the expression of nearby genes and immune pathways in MS risk.

Observational studies have shown that increased dietary intake of vitamin D is associated with reduced risk of MS, but it is unknown whether interventional vitamin D supplementation would yield protective effects on disease susceptibility (6, 27). Studies investigating the effects of vitamin D supplementation on MS progression have found associations between higher serum levels of vitamin D and reduced hazard of relapse and a reduction in incidence of T2 lesions (28, 29). Our results show how genetic variation in the vitamin D pathway affects MS disease risk. Future studies investigating the protective effects of vitamin D supplementation on disease risk and progression should consider genetic variation in vitamin D pathway genes as part of their design.

This study had several strengths. It is the first to investigate the association of VDR binding at a locus with MS among cases and controls. An investigation of variation in VDR binding among cases and controls would not be feasible in large sample sizes without utilizing MR methodology. We used previously identified VDR-BVs to create GIVs for VDR binding at a locus (19). Another strength was the large sample size and use of data from independent studies. There were also some limitations, including the possibility of horizontal pleiotropy,

biasing associations between GIV_{VDR} and MS. Each GIV_{VDR} was constructed using only one VDR-BV, preventing the use of standard sensitivity analyses for assessing the validity of MR assumptions in our VDR-BV analyses. However, our use of GIVs for 25(OH)D serum level in the analyses allowed us to estimate the association between GIV_{VDR} and MS unbiased by horizontal pleiotropy. Additionally, all but two of the VDR-BV were independent of known MS GWAS variant risks. MS cases and controls in the study were of European ancestry to reduce the impact of population stratification; however, this approach limits the generalizability of our findings to non-European populations. Further, VDR-BVs were identified in LCLs which are EBV-transformed B-cell lines. Variation in VDR binding associated with MS likely occurs in other lymphocytes, including CD4+ and CD8+ T cells and it is unknown whether the VDR-BVs active in LCLs are also active in these T cells. We were also only able to assess variation in VDR binding at 112 loci, which represent a small portion of genomic regions where VDR binds. Lastly, there are likely to be mechanisms other than VDR binding that underlie the observed VDR-MS associations reported here, including alternative splicing, effect of microRNAs, and combinatorial protein binding. Our study was designed to identify causal mechanisms of MS susceptibility that are mediated by altered VDR binding due to DNA variation.

This study is the first to demonstrate that genetic variation in VDR binding at a single locus contributes to MS susceptibility. Our results highlight the importance of the vitamin D pathway in MS pathogenesis. Our results are also relevant to other autoimmune and inflammatory diseases and several cancers for which a role for vitamin D has been suggested. Future studies of VDR binding and MS should identify VDR-BVs in lymphocytes not captured by LCLs, including CD4+ and CD8+ T cells.

Materials and Methods

MS case-control studies

Individual-level case-control data used in this research was from one USA-based case-control study and two Sweden-based MS case-control studies. MS cases and controls from the US were from the Kaiser Permanente Northern California (KPNC) MS Research Program (30). Additional KPNC controls were participants of the Genetic Epidemiology Research on Adult Health and Aging (GERA; dbGaP phs000674.v2. p2) (31). From Sweden, MS cases and controls were from the Nationwide Epidemiological Investigation of Multiple Sclerosis (EIMS), the Genes and Environment in Multiple Sclerosis (GEMS), IMSE study of the effect of immunomodulatory drugs, and the Karolinska Hospital STOPMS study (32, 33). Methods for identification and confirmation of MS diagnosis in each study are provided within the corresponding publications. Briefly, MS cases from KPNC, EIMS, GEMS, IMSE, and STOPMS had their disease status confirmed by an independent neurologist. All study participants provided written informed consent, and all studies obtained approval from the Institutional Review Boards of KPNC, local Ethical Committees, and the University of California, Berkeley.

UK Biobank

The UK Biobank (UKB, http://www.ukbiobank.ac.uk) is a prospective cohort study of approximately 500,000 individuals from the United Kingdom (34). Recruitment took place between 2006-2010 in 22 assessment centers located across the United Kingdom. Participants were aged 40-69 at the time of recruitment. MS cases were identified using UKB field 31043-0.0 "Source of report of G35 (multiple sclerosis)". Controls were defined as those not reporting MS symptoms or other demyelinating diseases at the time of enrollment and were frequency matched to cases by year of birth (±2 years) and sex at a ratio of 10 controls per MS case. Participants included in analyses were unrelated individuals of European ancestry with imputed genetic data and no sex chromosome aneuploidies. Those who had withdrawn consent or were recommended for genetic analysis exclusion were excluded. Data were accessed under the approval of UKB within project 69668. All participants gave prior written informed consent, and the study was conducted following the principles of the Declaration of Helsinki.

Genotype and exposure assessment

All participants in the MS case-control studies (KPNC, GEMS, EIMS, IMSE, and STOP-MS) completed an interview or self-reported questionnaire related to MS disease events, reproductive history, and environmental exposures (30–33). Participants provided blood or saliva samples for genotyping. SNP genotyping was performed using the Illumina Infinium 660K BeadChip Array and Human Omni Express Array (KPNC), Illumina Global Screening Array or Human Omni Expression (GEMS, EIMS, IMSE, and STOPMS), and Affymetrix Axiom Array (GERA). The data from GEMS, EIMS, IMSE, and STOPMS were analyzed in two cohorts depending on which genotyping array was used: Human Omni Express (OMNI) or Global screening array (GSA). Details of genotyping and imputation for KPNC (35), GSA AND OMNI (36), and UKB have been described previously. Participants with missing genotypes that met QC thresholds (info score>0.8, missingness per SNP<0.05, missingness per cohort<0.05, and minor allele frequency (MAF)>0.05) were imputed using the mean MAF within each study.

Genetic instruments for VDR binding

Mendelian Randomization (MR) is a method that uses one or more genetic variants associated with an exposure or other variable of interest as a genetic instrumental variable (GIV) to estimate causal associations between that exposure and an outcome. GIVs are calculated by summing the weighted contributions of individual SNPs, based on their effect sizes derived from GWASs or other genetic studies. In this study $GIV_{VDR\text{-}BV}$ for each individual were constructed using the following equation:

$$GIV_{VDR\text{-}BV_{ij}} = dosage_{ij} \times \beta_{VDR\text{-}BV_j}$$

where, $dosage_{ij}$ is the number of effect alleles at $VDR\text{-}BV_j$ carried by individual i and $\beta_{VDR\text{-}BV_i}$ is the effect size for the association between VDR-binding at a locus and alleles at $VDR\text{-}BV_j$. There are three primary MR assumptions: the variant(s) must be associated with the exposure, there are no unmeasured confounders of the association between the

genetic variant(s) and the outcome, and the genetic variants only affect the outcome through the exposure (8). The occupancy of VDR at a given genomic locus was the exposure of interest in this study. Our objective was to use SNPs associated with VDR binding at a locus previously identified by Gallone et al (37) as GIVs to identify associations between variation in VDR binding at a locus and MS susceptibility. Occupancy of VDR was quantified genome-wide at heterozygous sites using ChIP-exo sequencing read data from sixteen genotyped calcitriol-stimulated LCLs. To satisfy the first assumption of MR, at least one variant statistically associated with altered VDR occupancy at a locus was required (8). SNPs associated with VDR binding in cis were identified by observing differential VDR occupancy over each allele of a given heterozygous SNP. Allele-specific VDR binding was identified using AlleleSeq, modelling reads mapping to each allele using the binomial distribution and applying a two-tailed test (38). VDR-BVs were called at heterozygous sites with five or more ChIP-Exo reads, with at least read being mapped to each allele in order to reduce unwanted effects of PCR duplication. After accounting for copy number variation (32) and correction for multiple testing using AlleleSeq's computational simulation approach (False discovery rate < 2%), 305 VDR-BVs were identified. Of these, 112 VDR-BVs were present in all imputed case-control datasets for analysis. SNP-VDR binding effect sizes were estimated by regressing the read counts from ChIP-exo data at a heterozygous VDR binding region described above against the corresponding VDR-BV alleles. For each VDR-BV the GIV was derived by multiplying the SNP-VDR binding effect size by the number alleles carried by each participant (hereafter, GIV_{VDR}).

Pleiotropy and Bioavailability of Vitamin D

The causal effect estimates from MR analysis are unbiased in the absence of pleiotropic effects. A common approach to estimating the pleiotropic bias is based on over-dispersion across multiple instrumental SNPs for a given exposure. However, sensitivity analyses for multi-SNP MR instruments were not applicable as there was only one instrumental SNP for each GIV_{VDR}. Given that VDR binding is dependent upon the bioavailability of vitamin D, we used variants and summary statistics from two recent GWAS on serum 25(OH)D to construct GIVs for bioavailability of vitamin D (see below) to perform subgroup analyses (12, 13). Interaction between a GIV_{VDR} and bioavailability of 25(OH)D was considered evidence of an association between VDR binding and MS susceptibility not biased by horizontal pleiotropy.

Genetic instruments for 25(OH)D

SNPs and the estimated effect sizes from two recent GWAS on serum 25(OH)D were used to construct each GIV_{25OHD} (12, 13). GWAS summary statistics were extracted from MR-Base R platform (39). The extract_instruments() function in the TwoSampleMR R package was used to identify independent variants with GWAS $p < 1 \times 10^{-8}$ within genomic windows of 10,000Kb ($R^2 < 0.001$) using a European LD reference panel from 1000 Genomes Phase 3. These procedures yielded 8 SNPs for the Jiang et al. $GIV_{25(OH)D}$ and 103 SNPs for the Revez et al. $GIV_{25(OH)D}$ (SI Appendix, Tables S3 and S4). After identifying these SNPs, each $GIV_{25(OH)D}$ was calculated in Plink1.9 using --score (40) which uses the following equation:

$$GIV_{25(OH)D_i} = \sum_{k}^{K} \beta_k \times dosage_{ik}$$

where, $dosage_{ik}$ is the number of effect alleles at GWAS variant k carried by individual i and β_k is the effect size for the association between reduced 25(OH)D serum levels at and GWAS variant k.

Statistical methods

Logistic regression was used to estimate the associations between all GIVs (GIV_{25OHD} and GIV_{VDR}) and MS susceptibility within each study:

$$logit(P(MS = 1)) \sim \beta_0 + \beta_1 GIV_{25OHD}$$
$$logit(P(MS = 1)) \sim \beta_0 + \beta_1 GIV_{VDR_j}$$

Additionally, to estimate an association not biased by horizontal pleiotropy, we estimated interaction between each GIV_{VDR} and GIV_{25OHD} for MS susceptibility:

$$logit(P(MS = 1)) \sim \beta_0 + \beta_1 GIV_{VDR_j} + \beta_2 GIV_{25OHD} + \beta_3 \left(GIV_{VDR_j} \times GIV_{25OHD}\right),$$

where β_3 is the estimate of multiplicative interaction between VDR binding at a locus and the bioavailability of 25(OH)D. A significantly non-zero value of β_3 indicates that pleiotropy does not explain all the observed association and therefore, the association between VDR-BV and MS is causal.

All models were adjusted for sex (male or female), quintiles of birth year, carriage of the *HLA-DRB1*15:01* allele, and the first six genome-wide principal components. Random effects meta-analysis was used to combine study-specific associations. All meta-analyses were performed using the metagen() function in the Meta R package (41). The DerSimonian-Laird estimator was used to estimate the between-study variance. Between study heterogeneity was assessed using Cochran's Q statistic and Higgins & Thompson's I^2 statistic. p-values from associations between each GIV_{25OHD} and MS and the interaction parameter between each GIV_{VDR} and GIV_{25OHD} were corrected for multiple tests using the Benjamini-Hochberg method (42).

Annotation of VDR-BVs to the nearest gene transcription start sites (TSS) and GO enrichment analysis was performed using rGREAT (43). Tissue-specific expression quantitative trait loci (eQTL) and splicing quantitative trait loci (sQTL) for VDR-BVs were from the GTEx v8 Project (20). LD between VDR-BVs and MS GWAS risk variants was estimated using the ld_matrix() function from the *ieugwasr* R package using the European LD reference panel from 1000 Genomes Phase 3 (44).

All statistical analyses were conducted using R 4.0.2 (45).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Data Availability and Materials

Metadata for KPNC (https://doi.org/10.6084/m9.figshare.22145966.v1) as well as GSA and OMNI participants (https://doi.org/10.6084/m9.figshare.22047314.v1) is available. Complete KPNC data can be accessed by contacting Lisa Barcellos (lbarcellos@berkeley.edu) and Lynn Hollyer (lhollyer@berkeley.edu). Requests will be reviewed by the IRB, and data can be shared upon approval. Complete GSA and OMNI data can be accessed by contacting Ingrid Kockum (Ingrid.Kockum@ki.se). Genotype and phenotype data for UKB participants can be accessed from UKB (https://www.ukbiobank.ac.uk/enable-your-research/about-our-data/genetic-data). GERA data are available on dbGaP (phs000674.v2.p2). R scripts for regression analyses and meta-analyses are available here: https://doi.org/10.5281/zenodo.7647412. Summary statistics from Gallone et al. for VDR-BVs used in this study (SI Appendix, Table S5) and study-specific and meta-analysis associations between GIV_{VDR-BV} and MS susceptibility for all VDR-BVs are available in the supplement (SI Appendix, Tables S6-S9).

References

1. Rosati G. The prevalence of multiple sclerosis in the world: an update. Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology. 2001; 22: 117–39. [PubMed: 11603614]

2. Goldberg P. Multiple sclerosis: vitamin D and calcium as environmental determinants of prevalence. International Journal of Environmental Studies. 1974; 6: 19–27.

3. Munger KL, Levin LI, Hollis BW, Howard NS, Ascherio A. Serum 25-Hydroxyvitamin D Levels and Risk of Multiple Sclerosis. JAMA. 2006; 296: 2832. [PubMed: 17179460]

4. van der Mei IAF, et al. Past exposure to sun, skin phenotype, and risk of multiple sclerosis: case-control study. BMJ. 2003; 327: 316. [PubMed: 12907484]

5. Kampman MT, Wilsgaard T, Mellgren SI. Outdoor activities and diet in childhood and adolescence relate to MS risk above the Arctic Circle. Journal of Neurology. 2007; 254: 471–477. [PubMed: 17377831]

6. Munger KL, et al. Vitamin D intake and incidence of multiple sclerosis. Neurology. 2004; 62: 60–5. [PubMed: 14718698]

7. Ascherio A, Munger KL, Simon KC. Vitamin D and multiple sclerosis. The Lancet Neurology. 2010; 9: 599–612. [PubMed: 20494325]

8. Smith GD, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003; 32: 1–22. [PubMed: 12689998]

9. Rhead B, et al. Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk. Neurology Genetics. 2016; 2: e97. [PubMed: 27652346]

10. Mokry LE, et al. Vitamin D and Risk of Multiple Sclerosis: A Mendelian Randomization Study. PLOS Medicine. 2015; 12 e1001866 [PubMed: 26305103]

11. Gianfrancesco MA, et al. Evidence for a causal relationship between low vitamin D, high BMI, and pediatric-onset MS. Neurology. 2017; 88: 1623–1629. [PubMed: 28356466]

12. Revez JA, et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. Nature Communications. 2020; 11: 1–12.

13. Jiang X, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nature Communications. 2018; 9: 260.

14. Bikle DD. Vitamin D metabolism, mechanism of action, and clinical applications. Chemistry & biology. 2014; 21: 319–29. [PubMed: 24529992]

15. Ramagopalan SV, et al. A ChIP-seq defined genome-wide map of vitamin D receptor binding: Associations with disease and evolution. Genome Research. 2010; 20: 1352–1360. [PubMed: 20736230]

16. Tuoresmäki P, Väisänen S, Neme A, Heikkinen S, Carlberg C. Patterns of Genome-Wide VDR Locations. PLoS ONE. 2014; 9 e96105 [PubMed: 24787735]

17. Pike JW, Meyer MB. The Vitamin D Receptor: New Paradigms for the Regulation of Gene Expression by 1,25-Dihydroxyvitamin D3. Endocrinol Metab Clin North Am. 2010; 39: 255–269. [PubMed: 20511050]

18. Wöbke TK, Sorg BL, Steinhilber D. Vitamin D in inflammatory diseases. Frontiers in Physiology. 2014; 5: 244. [PubMed: 25071589]

19. Gallone G, et al. Identification of genetic variants affecting vitamin D receptor binding and associations with autoimmune disease. Human Molecular Genetics. 2017; 26: 2164–2176. [PubMed: 28335003]

20. GTEx Portal. 2023 May 18.

21. Saeki K. The B cell-specific major raft protein, Raftlin, is necessary for the integrity of lipid raft and BCR signal transduction. The EMBO Journal. 2003; 22: 3015–3026. [PubMed: 12805216]

22. Gauderman WJ, et al. Update on the State of the Science for Analytical Methods for Gene-Environment Interactions. American Journal of Epidemiology. 2017; 186: 762–770. [PubMed: 28978192]

23. Watanabe A, et al. Raftlin Is Involved in the Nucleocapture Complex to Induce Poly(I:C)-mediated TLR3 Activation. J Biol Chem. 2011; 286: 10702–10711. [PubMed: 21266579]

24. Tatematsu M, et al. Raftlin Controls Lipopolysaccharide-Induced TLR4 Internalization and TICAM-1 Signaling in a Cell Type–Specific Manner. The Journal of Immunology. 2016; 196: 3865–3876. [PubMed: 27022195]

25. Zheng C, Chen J, Chu F, Zhu J, Jin T. Inflammatory Role of TLR-MyD88 Signaling in Multiple Sclerosis. Frontiers in Molecular Neuroscience. 2020; 12

26. Hendrickx DAE, et al. Gene Expression Profiling of Multiple Sclerosis Pathology Identifies Early Patterns of Demyelination Surrounding Chronic Active Lesions. Frontiers in Immunology. 2017; 8

27. Sintzel MB, Rametta M, Reder AT. Vitamin D and Multiple Sclerosis: A Comprehensive Review. Neurology and therapy. 2018; 7: 59–85. [PubMed: 29243029]

28. Simpson S, et al. Higher 25-hydroxyvitamin D is associated with lower relapse risk in multiple sclerosis. Ann Neurol. 2010; 68: 193–203. [PubMed: 20695012]

29. Mowry EM, et al. Vitamin D status predicts new brain magnetic resonance imaging activity in multiple sclerosis. Ann Neurol. 2012; 72: 234–240. [PubMed: 22926855]

30. Briggs FBS, et al. Adverse socioeconomic position during the life course is associated with multiple sclerosis. J Epidemiol Community Health. 2014; 68: 622–629. [PubMed: 24577137]

31. Banda Y, et al. Characterizing race/ethnicity and genetic ancestry for 100,000 subjects in the genetic epidemiology research on adult health and aging (GERA) cohort. Genetics. 2015; 200: 1285–1295. [PubMed: 26092716]

32. Hedström AK, Hillert J, Olsson T, Alfredsson L. Alcohol as a Modifiable Lifestyle Factor Affecting Multiple Sclerosis Risk. JAMA Neurology. 2014; 71: 300. [PubMed: 24395432]

33. Hedström A, Åkerstedt T, Olsson T, Alfredsson L. Shift work influences multiple sclerosis risk. Multiple Sclerosis Journal. 2015; 21: 1195–1199. [PubMed: 25698167]

34. Sudlow C, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Medicine. 2015; 12 e1001779 [PubMed: 25826379]

35. Rhead B, et al. Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk. Neurology Genetics. 2016; 2: e97. [PubMed: 27652346]

36. Baranzini S, Sawcer S, International Multiple Sclerosis Genetics Consortium, MultipleMS Consortium, *. Genetic analysis of multiple sclerosis severity identifies a novel locus and implicates CNS resilience as a major determinant of outcome. 2022; doi: 10.21203/rs.3.rs-1723574/v1

37. Gallone G, et al. Identification of genetic variants affecting vitamin D receptor binding and associations with autoimmune disease. Human Molecular Genetics. 2017; 26: 2164–2176. [PubMed: 28335003]

38. Rozowsky J, et al. AlleleSeq: analysis of allele-specific expression and binding in a network framework. Mol Syst Biol. 2011; 7: 522. [PubMed: 21811232]

39. Hemani G, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018; 7 e34408 [PubMed: 29846171]

40. Chang CC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015; 4: 7. [PubMed: 25722852]

41. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health. 2019; 22: 153–160. [PubMed: 31563865]

42. Yekutieli D, Benjamini Y. The control of the false discovery rate in multiple testing under dependency. Ann Statist. 2001; 29: 1165–1188.

43. Gu, Z. rGREAT: Client for GREAT Analysis. 2021.

44. Hemani, G. ieugwasr: R interface to the IEU GWAS database API. 2021.

45. R Core Team. — European Environment Agency. 2021. October 1, 2020

Significance Statement

There is evidence for a causal association between 25-hydroxyvitamin D (25(OH)D) serum levels and multiple sclerosis (MS), but the underlying mechanisms are unknown. We conducted a Mendelian Randomization study to estimate causal associations between vitamin D receptor (VDR) binding at a locus and MS susceptibility using genetic instrumental variables constructed from variants previously associated with allele-specific VDR-binding (VDR-binding variant; VDR-BV) and serum 25(OH)D. Two VDR-BV (rs2881514; rs2531804) were independently associated with MS and rs2881514 demonstrated significant interaction with 25(OH)D, providing evidence for a causal association between rs2881514 and MS. Our results demonstrate that genetic variation in individual VDR binding sites contributes to MS and are relevant to other autoimmune diseases for which a role for vitamin D has been implicated.
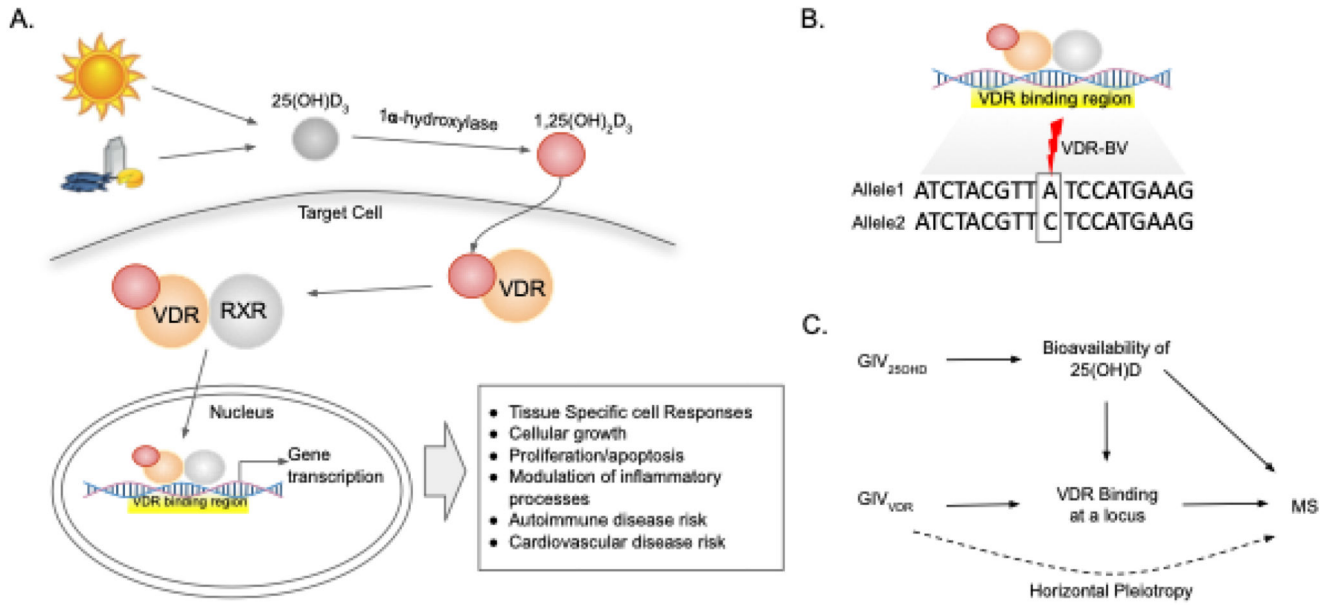
Figure 1.
A. Representation of the vitamin D pathway. B. Example of a vitamin D receptor binding variant (VDR-BV) location within a VDR binding region. VDR-BVs are single nucleotide polymorphisms associated with increased or decreased VDR binding at a VDR binding region. C. Directed acyclic graph depicting our mendelian randomization analyses. Genetic variants associated with altered VDR binding at a locus (GIV_{VDR-BV}) are used to estimate association between VDR-Binding at a locus and MS susceptibility within our MS case-control datasets. Genetic variants from genome-wide association studies on serum vitamin D levels ($GIV_{25(OH)D}$) are used to estimate the association between VDR binding and MS unbiased by horizontal pleiotropy (dashed-line).
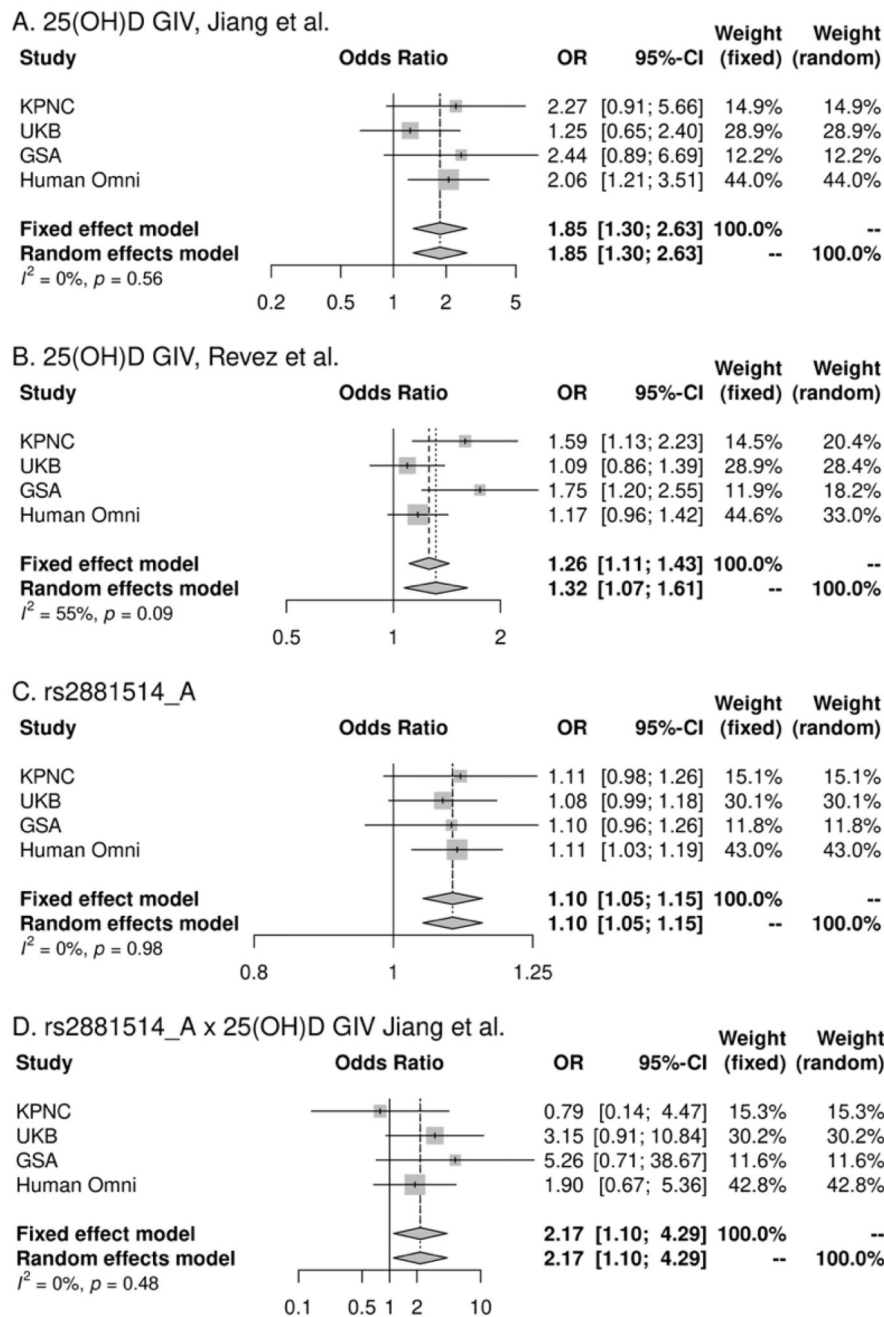
Figure 2.

Results from meta-analyses of association between 25(OH)D genetic instrumental variables (GIV$_{25OHD}$) and VDR-BV rs2881514 and multiple sclerosis. GIVs$_{25OHD}$ calculated using summary statistics from genome-wide association studies on serum 25(OH)D levels. A) GIVs$_{25OHD}$ calculated using summary statistics from Jiang et al. 2019 (13); B) GIVs$_{25OHD}$ calculated using summary statistics from Revez et al. 2020 (12). (C) Meta-analysis results for association between GIV$_{25(OH)D}$ rs2881514 and multiple sclerosis and (D) Evidence of interaction between rs2881514 and Jiang et al. 25(OH)D genetic instrumental variable. $\tau^2=$

estimate of between study heterogeneity; I^2=proportion of variance due to heterogeneity; and p=p-value for Cochran's Q test of heterogeneity.

Table 1

Characteristics of MS cases and controls.

	KPNC		UKB		GSA		OMNI	
	Case	Control	Case	Control	Case	Control	Case	Control
n	1084	10956	2087	20870	3718	1180	6709	5881
Sex, n (%)								
Male	215 (19.8)	2165 (19.8)	576 (27.6)	5760 (27.6)	1031 (27.7)	342 (29.0)	1808 (26.9)	1455 (24.7)
Female	869 (80.2)	8791 (80.2)	1511 (72.4)	15110 (72.4)	2687 (72.3)	838 (71.0)	4901 (73.1)	4426 (75.3)
Birth year, mean (SD)	1957.6 (8.9)	1957.1 (8.6)	1952.7 (7.6)	1952.7 (7.6)	1973.1 (12.4)	1968.9 (13.7)	1961.0 (13.7)	1960.5 (13.3)
*DRB1*15:01* carrier, n (%)								
0 alleles	506 (46.7)	8280 (75.6)	1050 (50.3)	15271 (73.2)	1621 (43.6)	843 (71.4)	2913 (43.4)	4052 (68.9)
1 or 2 alleles	578 (53.3)	2676 (24.4)	1037 (49.7)	5599 (26.8)	2097 (56.4)	336 (28.5)	3796 (56.6)	1829 (31.1)
25(OH)D GIV, Jiang et al.[a] (mean (SD))	0.285 (0.07)	0.288 (0.07)	0.292 (0.07)	0.293 (0.07)	0.289 (0.07)	0.293 (0.07)	0.286 (0.07)	0.290 (0.07)
25(OH)D GIV, Revez et al.[a] (mean (SD))	2.150 (0.12)	2.163 (0.19)	2.184 (0.19)	2.187 (0.19)	2.057 (0.19)	2.074 (0.18)	2.066 (0.19)	2.070 (0.19)

Abbreviations: GSA, Swedish cohort genotyped on global screening array; Human Omni, Swedish cohort genotyped on human omni express array; KPNC, Kaiser Permanente Northern California MS case-control study; GIV, Genetic Instrumental Variable; UKB, UK Biobank

[a]GIVs calculated from independent GWAS variants (Linkage disequilibrium $R^2<0.001$) with $p<5 \times 10^{-8}$.

Table 2

Vitamin D Receptor binding variants (VDR-BVs) instrumental variable (IV) associated with multiple sclerosis susceptibility at p<0.05.

GIV$_{VDR}$[a]	chr:bp	GREAT annotation[b]	OR[c]	95% CI	P	P$_{FDR}$[c]
rs2881514_A	chr3:16553679	RFTN1 (+1533); OXNAD1 (+246974)	1.10	1.05-1.15	9.4E-05	0.011
rs2531804_A	chr6:28411302	ZSCAN23 (-59)	0.82	0.73-0.92	6.4E-04	0.036
rs12048389_T	chr1:107538722	PRMT6 (-60578)	0.92	0.87-0.99	1.8E-02	0.660
rs55792977_T	chr2:65650863	SPRED2 (+8447); ACTR2 (+195893)	1.05	1.00-1.09	3.2E-02	0.660
rs7309003_C	chr12:97751789	NEDD1 (+450546)	0.95	0.91-1.00	3.2E-02	0.660
rs10995246_C	chr10:64391845	ADO (-172670); ZNF365 (+257895)	1.06	1.00-1.13	3.5E-02	0.660
rs10232857_C	chr7:55601335	VOPP1 (+38882); LANCL2 (+168195)	0.93	0.87-1.00	4.9E-02	0.780

Abbreviations: bp, base pair; chr; chromosome; CI, Confidence Interval; FDR, false discovery rate; GIV, Genetic Instrumental Variable; GREAT, Genomic Regions Enrichment of Annotations Tool; OR, Odds Ratio; VDR-BV, Vitamin D Receptor Binding Variant.

[a]VDR-BV using as instrumental variable for allelic specific binding. Allele included is the allele associated with increased binding.

[b]Distance (bp) from VDR-BV to TSS of the nearest upstream and downstream gene from rGREAT.

[c]Odds ratios are from random effect meta-analyses combining estimates from the four studies. ORs for association between VDR-BV and MS.