Relevance of Multiple Sclerosis Severity Genotype in Predicting Disease Course: A Real-World Cohort.

Karim L. Kreft, MD, PhD, MSc(Hons) [ID],[1] Emeka Uzochukwu, PhD,[2] Sam Loveless, PhD,[2] Mark Willis, MD, PhD,[1,2] Ray Wynford-Thomas, MD,[2] Katharine E. Harding, MD, PhD,[3] Peter Holmans, PhD,[2] Michael Lawton, PhD,[4] Emma C. Tallantyre, MD, PhD [ID],[1,2] and Neil P. Robertson, MBBS, MD[1,2]

Objective: Currently, 233 genetic loci are known to be associated with susceptibility to multiple sclerosis (MS). Two independent pivotal severity genome-wide association studies recently found the first genome-wide significant single-nucleotide variant (SNV; rs10191329[A]) and several other suggestive loci associated with overall disability outcomes. It is now important to understand if these findings can influence individual patient management.

Methods: We assessed whether these progression SNVs are associated with detailed clinical phenotypes in a well-characterized prospective cohort of 1,455 MS patients. We used logistic regression, survival analysis, and propensity score matching to predict relevant long-term clinical outcomes.

Results: We were unable to detect any association between rs10191329[A] and a range of clinically relevant outcomes (eg, time to Expanded Disability Status Scale milestones, age-related MS severity score, anatomical localization at onset or during subsequent relapses, annualized relapse rate). In addition, an extremes of outcome case–control analysis using a propensity score matching for genotype detected no association between disease severity and rs10191329[A]. However, we were able to replicate the association of two suggestive SNVs (rs7289446[G] and rs868824[C]) with the development of fixed disability, albeit with modest effect sizes, and the association of HLA-DRB1*1501 with age at onset.

Interpretation: Identification of rs10191329[A] and other suggestive SNVs are of considerable importance in understanding pathophysiological processes associated with MS severity. However, it is unlikely that individual genotyping can currently be used in a clinical setting to guide disease management. This study shows the importance of independent replication of genome-wide association studies associated with disease progression in neurodegenerative disorders.

ANN NEUROL 2023;00:1–12

Introduction

Multiple sclerosis (MS) is a chronic disease with an early inflammatory and later neurodegenerative components, of which the former is better understood. All currently licensed disease-modifying treatments (DMT) for MS are designed to limit neuroinflammation, and drugs that directly target neurodegenerative changes in MS are lacking. As a result, many people with MS (pwMS) still develop permanent disability, including the need for a walking aid or wheelchair.[1] A component of MS pathophysiology consists of heritable factors and, to date, all genome-wide association studies (GWAS) have addressed

disease susceptibility, uncovering a total of 233 risk single-nucleotide variants (SNVs).[2] Efforts to detect SNVs associated with disease severity have been unproductive, until recently[3-5] when the International Multiple Sclerosis Genetics Consortium (IMSGC) performed a pivotal severity GWAS,[6] while the MSBase consortium also published their independent GWAS of 1,813 pwMS of European descent.[7] The IMSGC used an extremes of outcome analysis in 22,000 pwMS to assess SNVs associated with the development of sustained disability (primary outcome was age-related multiple sclerosis severity score [ARMSS], and time to EDSS milestones was a secondary outcome). They found rs10191329[A] reached genome-wide significance for a faster progression toward confirmed disability progression, and also that rs149097173 was a suggestive SNV.[6] Interestingly, both SNVs have a presumed biological function within the central nervous system, making them probable candidate loci for neurodegeneration, whereas the known MS susceptibility SNVs are mainly associated with immunological functions. In the MSBase GWAS, rs10191329 was not significant; however, some alternative variants were associated with sustained disability, although none reached genome wide significance. Interestingly, all identified variants have a high expression within the central nervous system, mainly in the cerebellar cortex. We assessed whether these important findings could be replicated in a large, independent, longitudinal, prospective cohort of pwMS, and to explore whether a more detailed analysis of clinical outcomes might reveal clinical utility.

Materials and Methods

Participants

All consecutive MS patients with longitudinal follow-up data in this prospective cohort study embedded within the South Wales MS registry and with available genotyping were included in the present study. Exclusion criteria were non-Caucasian ethnicity, because all GWAS results have only been validated in pwMS of Caucasian origin and/or patients with a clinically isolated syndrome without evidence of further disease activity or progression to MS. All included patients fulfilled the 2017 McDonald Criteria for MS.[8] The South Wales MS registry was initiated in 1985, and longitudinal demographic characteristics and clinical data, including DMT, have been recorded in a standardized database with outcome parameters, such as imaging findings and longitudinal EDSS data, since that time. This study was approved by the Research Ethics Committee of Health and Care Research Wales, reference numbers 05/WSE/03/111, 19/WA/0289 and 19/WA/0058, and all participants provided written informed consent before inclusion. None of the participants of the South Wales MS registry were included in either the IMSGC or MSBase GWAS.

Genotyping

In all PwMS, DNA was extracted from blood in ethylene diamine tetra acetic acid (EDTA)-containing tubes (in some cases from saliva using an Isohelix collection kit), and stored for subsequent genotyping at −80°C. Genotyping was performed using either Immunochip or the Infinium CoreExome-24 v2 or v3 chip (both Illumina, San Diego, CA, USA) according to the manufacturers' instructions. A variant was retained if it had a call rate >0.95%, minor allele frequency (MAF) >0.01, and Hardy–Weinberg equilibrium $>1*10^{-6}$. Individuals were excluded if they were missing >5% of the variants, non-European, or heterozygosity <−0.1 or >0.1. Principal components were used to check for population stratification and relatedness coefficient >0.125. Imputation was performed for all samples and variants, which passed quality control steps, and imputation was conducted for chromosomes 1 to 22 using Minimac4, Haplotype Reference Consortium panel (r1.1), and Eagle v2.4 at the TOPMed Imputation Server. After imputation, variants with imputation quality of $R^2 < 0.3$ and MAF <1% were excluded. All cohorts were merged after quality control was completed, leaving ∼5 million variants. The SNV, rs10191329, was extracted. The call rate of rs10191329 was 98.42%. The first 10 principal components (principal component analysis) were calculated for variants with a MAF ≥0.05, missing data <0.5, and Hardy Weinberg equilibrium 1×10^{-10}. SNV were subsequently LD pruned with 1,000 kb and lastly the principal component analysis were calculated using PLINK 2.0 (www.cog-genomics.org/plink/2.0/).[9] For survival analysis (see below), comparisons were made between homozygous non-risk carriers (CC genotype) and heterozygous and homozygous risk carriers (CA and AA respectively) pooled together to increase statistical power, because of the small number of patients with a homozygous risk allele due to the low MAF in people of European descent. We also assessed whether the results differed when comparing the three different genotypes (Fig. S3).

EDSS and Derived Outcome Measures

All EDSS scores were recorded in a standardized form by trained physicians blinded to genomic data. In this study, we only included EDSS data when a patient was physically examined (all phone- or questionnaire-based EDSS scores were excluded from analysis). In keeping with the IMSGC severity GWAS, EDSS scores were subsequently converted to ARMSS.[10]

Annualized Relapse Rate

Relapses were recorded if a subacute neurological worsening occurred with a duration of at least 24 hours without evidence of any provocative factor, such as infections. The annualized relapse rate was calculated as the number of relapses divided by the total follow-up duration since onset of the disease. Patients with primary progressive MS were omitted for calculation of annualized relapse rate.

Disease-Modifying Treatment

The use of DMT was defined as either (1) never receiving a DMT, (2) the use of a moderate efficacy (ME) DMT

(interferons, glatiramer acetate, teriflunomide, and dimethyl-fumarate), (3) high efficacy (HE) DMT (fingolimod, siponimod, cladribine, natalizumab, ocrelizumab, ofatumumab, and alemtuzumab), or (4) both ME and HE DMT.[11]

Anatomical Localization According to the European Database for Multiple Sclerosis Criteria

Both the anatomical localization at onset of the disease and the relative frequency of anatomical localization within the central nervous system of all relapses were compared between risk carriers and carriers of at least one protective allele. Localization was determined according to symptom domains from the European Database for Multiple Sclerosis project, with an additional cerebellar domain.[12]

Statistical Analysis

The distribution of all linear outcomes was determined by visually inspecting QQ-plots and by calculating the Shapiro–Wilk test. The majority of data were not normally distributed and, therefore, all comparisons were made using nonparametric statistical tests, unless stated otherwise. For linear regression analysis, we applied a rank-based inverse normal transformation (RINT) to the ARMSS score. After RINT, the data were normally distributed (Fig. S2).

Linear regression was performed with RINT ARMSS as a dependent variable, and rs10191329[A] sex, age at onset, localization at onset, cumulative frequency of localization of relapses, annualized relapse rate, and the use of DMTs as independent variables (for specific details of the models applied, please refer to the manuscript). Similar linear regression models were constructed to validate the MSBase-suggestive SNVs (rs7289446, rs1207401, rs10967273, rs698805, rs295254, rs9643199, rs2776741, rs7070182, rs868824, and rs3135388), replacing rs10191329 as an independent predictor.

Survival analysis was applied to time-dependent variables. We used time to EDSS of 4 (maximum walking ability without aide of 500 meters), EDSS 6 (requiring a walking aide to walk

Table 1. Demographic characteristics

	rs10191329[CC] (n = 1,001)	rs10191329[CA] (n = 417)	rs10191329[AA] (n = 37)	p value
Mean age at onset, yr (SD)	33 (11)	33 (11)	34 (11)	0.80
Female sex	69%	71%	68%	0.80
Oligoclonal bands % total number of CSF performed (CSF)	83% (512)	83% (204)	95% (18)	0.50
Patients with primary progressive MS (% of total included patients)	102 (10%)	40 (9.9%)	4 (11%)	0.90
No. HLA-DRB1*1501 (% of total)				
Homozygous negative	497 (50%)	181 (42.7%)	13 (35%)	
Heterozygous	444 (44%)	203 (48.7%)	18 (49%)	**0.024**
Homozygous positive	60 (6.0%)	33 (8.6%)	6 (16%)	
Median annualized relapse rate[a]	0.022 0.012–0.039)	0.024 (0.012–0.037)	0.030 (0.016–0.038)	0.70
Mean disease duration to last EDSS measurement (SD, in years)	14 (12)	14 (12)	13 (12)	>0.90
Exposure to disease-modifying treatments (% of total per genotype)				
Never	734 (73.3%)	301 (72.2)	30 (81.1%)	
Only moderate efficacy	175 (17.5%)	70 (16.8)	5 (13.5%)	0.58[b]
Only high efficacy	38 (3.8)	25 (6.0)	1 (2.7%)	
Switch between moderate and high efficacy	54 (5.4%)	21 (5.0)	1 (2.7%)	

Note: All data represents the median with interquartile range, unless stated otherwise.
[a]Relapsing onset patients only.
[b]Pearson's χ^2-test with simulated p value (based on 2,000 replicates).
Abbreviations: CSF = cerebrospinal fluid; EDSS = Expanded Disability Status Score.

for 100 meters), EDSS 8 (wheelchair dependent), and the time to reach secondary progressive MS (relapsing onset patients only) as relevant outcome measures for disease severity. Survival curves using the Kaplan–Meier approach and Cox proportional hazards models were constructed using sex and age at onset as covariates. Cox proportional hazards models using identical covariates from the linear regression models showed similar results (Table S2).

We calculated a weighted genomic risk score for MS susceptibility (wGRS; of 181 genome-wide significant SNV, the remaining 20 MS susceptibility SNVs were not available to study after quality controls and imputation) and a human leukocyte antigen (HLA) genomic burden score (HLAGB; of 10 SNV), as previously described.[13] In each score, allele counts were weighted by the corresponding effect size from previous published GWAS.[2] Next, we applied generalized linear modeling RINT ARMSS score as the dependent variable, and sex, age at last EDSS, and either (1) wGRS, (2) HLAGB, or (3) combined wGRS and HLAGB as predictors in the entire cohort of 1,455 MS patients. In addition, we calculated an MS severity wGRS-based 1 genome-wide significant SNV and 4 suggestive ($P < 5 \times 10^{-6}$) SNVs of the IMSGC severity GWAS (rs10191329, rs61215450, rs2876767, rs4251626, and rs194722 were included).[6] For all polygenic scores, we used all SNVs available after imputation, the remaining SNVs were not present in our genome-wide data. We used Pearson's correlation coefficient to

assess a relationship with ARMSS, and age at onset and MS severity wGRS. The *t* test was used to compare MS severity wGRS and type of disease onset. Finally, we performed a survival analysis comparing the highest and lowest quantile of the MS severity PRS, and the time to EDSS milestones (as described above).

In the final analysis, we compared homozygous risk carriers with homozygous non-risk carriers. We excluded 2 homozygous risk carriers due to missing data for matching. We applied propensity score matching without replacement with a nearest neighbor approach. The case-to-control ratio was set at 1:4 to optimize statistical power[14] using logistic regression with age at onset, sex, relapsing versus progressive onset, and the use of DMT (classified as never on DMT, ME drugs only, HE drugs only, or switching from moderate to HE DMT)[11] throughout the disease course as covariates. After matching, all standardized mean differences for the covariates were <0.03, indicating adequate balance between the groups (Fig. S4). Subsequently, ARMSS scores were compared between 35 homozygous risk carriers and 140 homozygous non-risk carriers, and survival analysis was applied for time to EDSS 4, 6, and 8 and secondary progression.

All analyses were performed using R studio (version 2022.02.3, build 492, and R version 3.6.3; The R Foundation for Statistical Computing, Vienna, Austria) using the packages
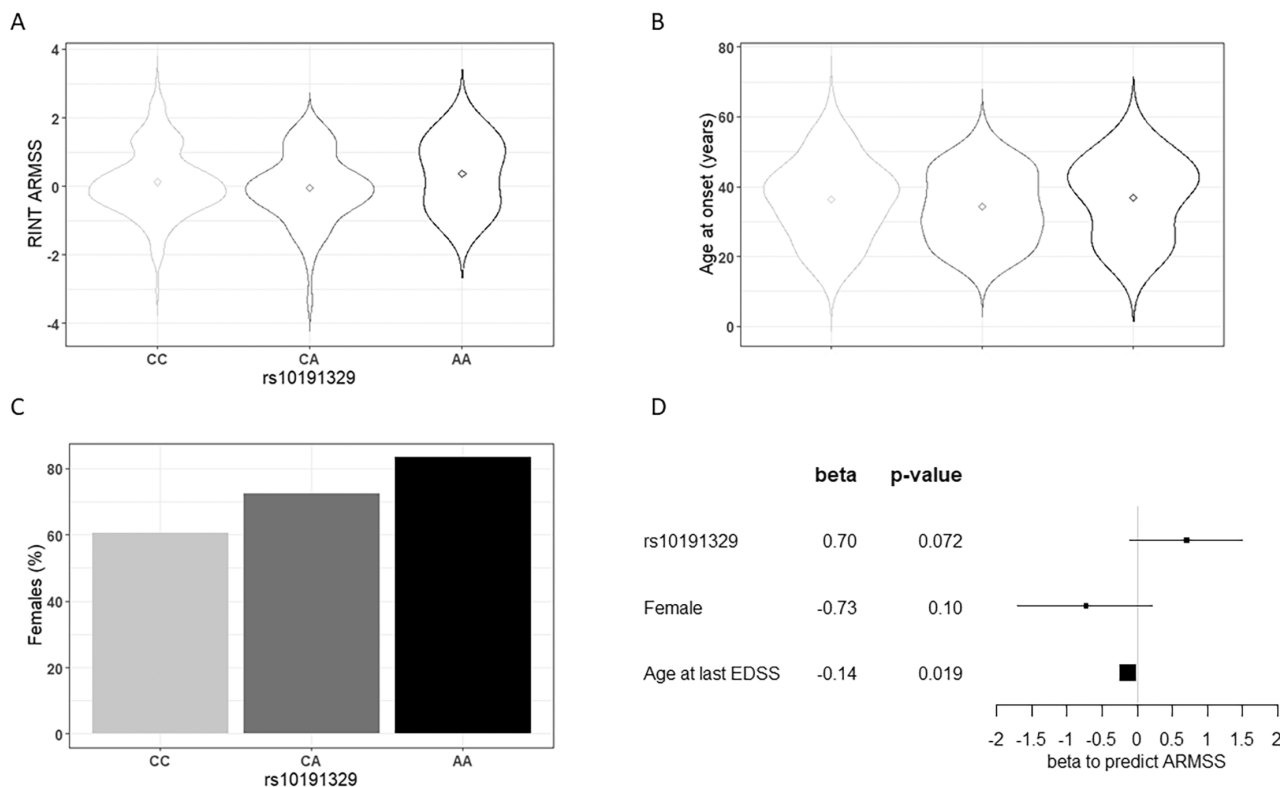

Figure 1: rs10191329[AA] **is not associated with an increased risk of developing more disability in 277 elderly multiple sclerosis patients with longstanding disease.** (A) Rank-based inverse normalized (RINT) age-related multiple sclerosis severity score (ARMSS) scores based on last known expanded disability status score (EDSS), stratified according to rs10191329. (B) Age at onset, stratified according to rs10191329. (C) Female-to-male ratio, stratified according to rs10191329. (D) Forest plot of regression model to predict ARMSS based on rs10191329 with relevant covariates (similar to the International Multiple Sclerosis Genetics Consortium study).

plyr (1.8.7), tidyverse (1.3.1), survival (3.2.11), survminer (0.4.9), janitor (2.1.0), gtsummary (1.6.0), rstatix (0.7.0), broom (0.8.0), readr (1.4.0), and naniar (0.6.1). Rank-based inverse normalization was calculated using RNOmni (1.0.0). ARMSS scores were calculated using the ms.sev package (1.0.4). Propensity score matching was performed using MatchIt (4.4.0) and optmatch (0.10.5). Graphs were also constructed in R studio using packages ggplot2 (from tidyverse package), patchwork (1.1.1), ggplotify (0.1.0), plotly (4.10.0), ggsignif (0.6.3), ggpubr (0.4.0), forestplot (2.0.1), cowplot (1.1.1), and ggimage (0.3.1). In the present study, we used a candidate gene approach and therefore P-values <0.05 without adjustment for multiple testing were considered statistically significant.

Results

rs10191329[A] Not Predicting Age-Related MS Severity Score

From the South Wales MS registry, a prospective cohort study following MS patients from 1985 onwards, 1,534 patients were genotyped, from which 1,455 were included in the current study (Table 1 and Fig. S1 for details on exclusion). Baseline characteristics were similar between rs10191329 risk and non-risk allele carriers, although a slightly higher percentage of HLA-DRB1*1501 carriers was observed in rs10191329 risk allele carriers ($p = 0.02$). The frequency of rs10191329 AA carriers in our study was 2.5% and CA carriers approximately 28.6%, which is comparable with the IMSGC GWAS findings.[6] From our total cohort, we first selected all pwMS who fulfilled criteria for inclusion in the IMSGC study (elderly age with a longstanding diagnosis of MS on the date of EDSS measurement). A total of 277 pwMS were included in this analysis. No significant differences were found between rs10191329 AA, CA, and CC carriers in the last known ARMSS score (Fig. 1A), the age at onset (Fig. 1B), or sex (Fig. 1C).

Next, we constructed a linear regression model, using rank-based inverse normalized ARMSS score as a
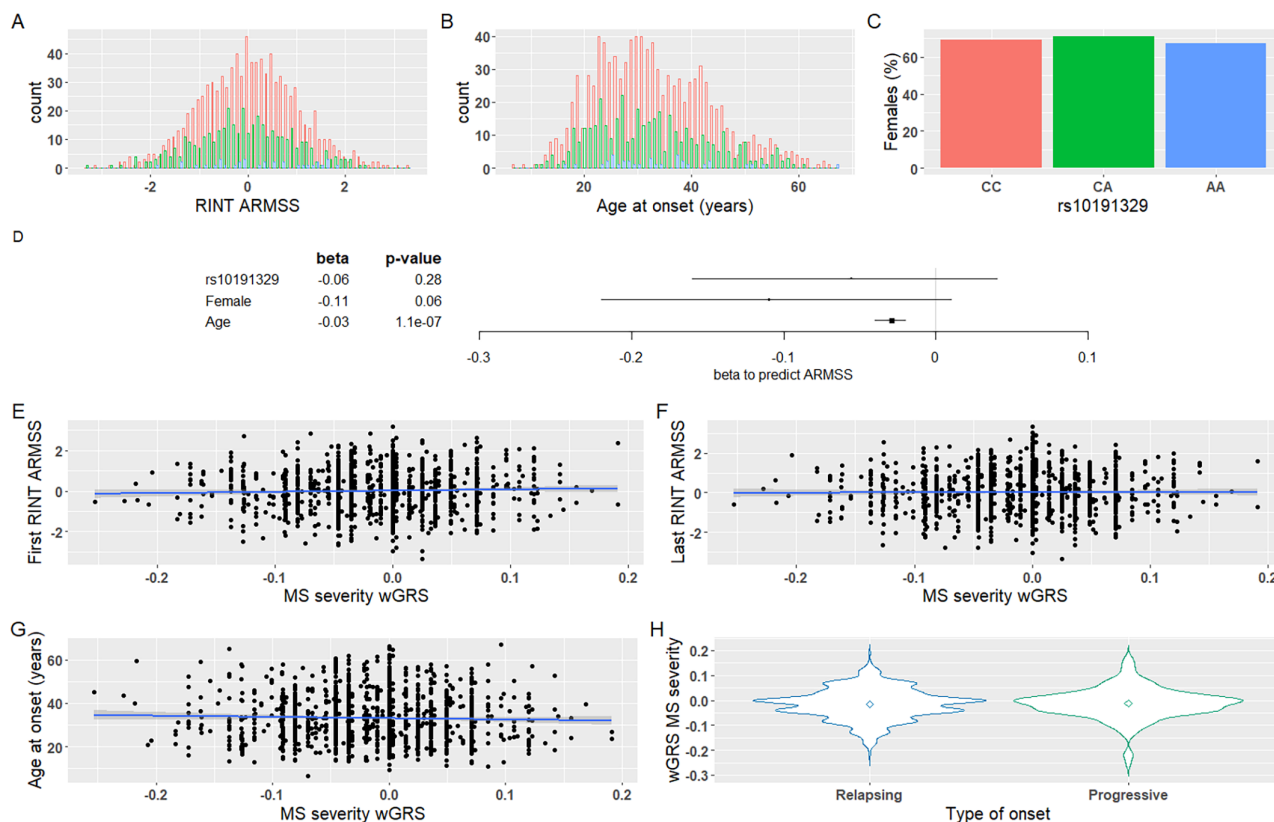

Figure 2: rs10191329[A] and multiple sclerosis (MS) severity-weighted genomic risk score are not associated with an increased risk to develop more disability in 1,455 MS patients. Stratified analysis according to rs10191329 carriership for (A) rank-based inverse normalized (RINT) age-related multiple sclerosis severity score (ARMSS) based on the last known Expanded Disability Status Score (EDSS). (B) Age at onset. (C) Sex ratio. (D) Forest plot of regression model to predict age-related multiple sclerosis severity score (ARMSS) based on rs10191329 with relevant covariates. Red represents rs10191329[CC], green rs10191329[CA], and blue rs10191329[AA]. Correlation between MS severity-weighted genomic risk score (wGRS) and (E) first ARMSS, (F) last ARMSS, and (G) age at onset. (H) No significant differences in the MS severity wGRS were observed between patients with a relapsing and a progressive onset of MS.
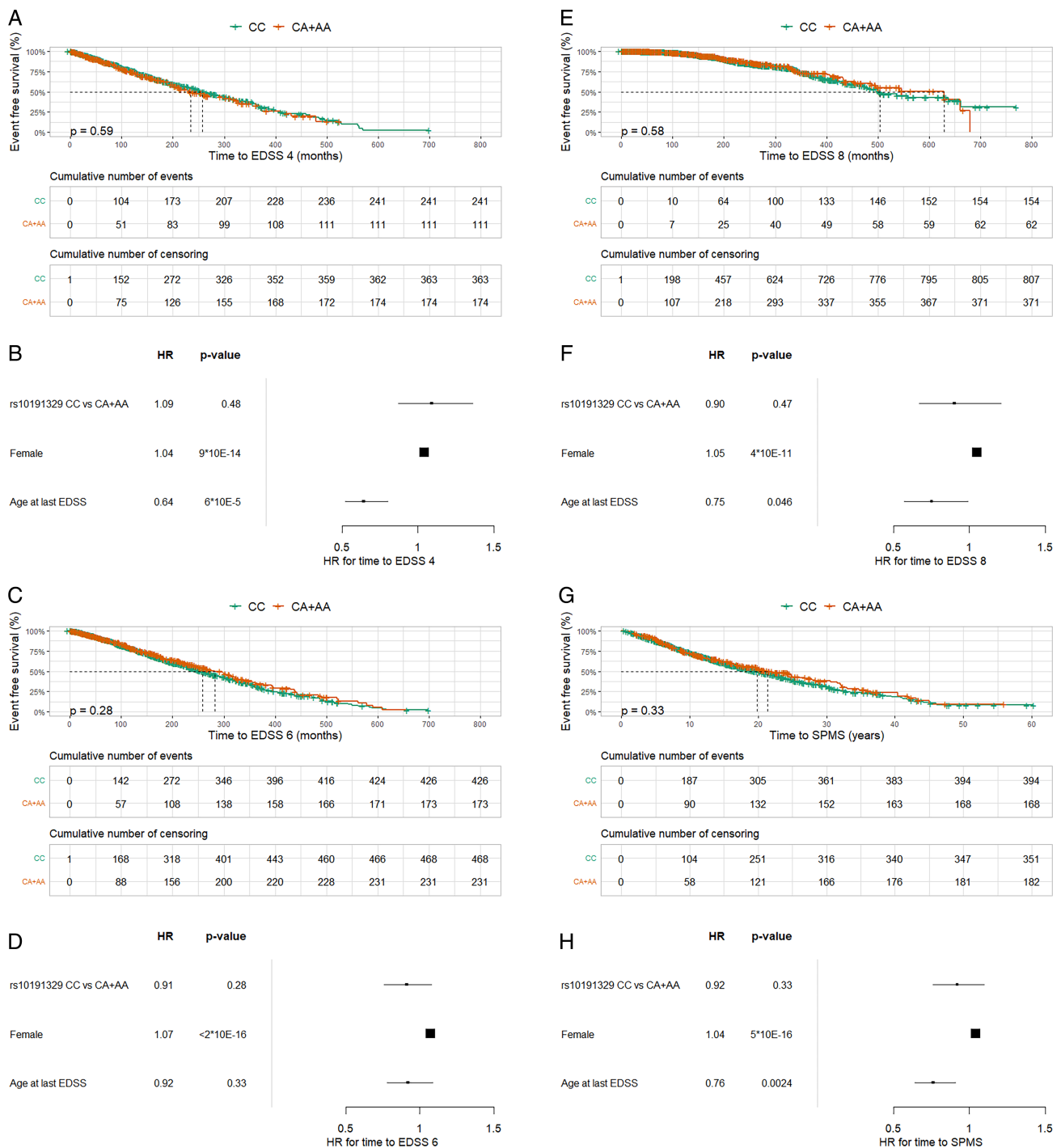
Figure 3: rs10191329 is not associated with time to reach expanded disability status score (EDSS) milestones or time to secondary progressive multiple sclerosis. Kaplan–Meier survival curves for time to (A) Expanded Disability Status Score (EDSS) 4, (C) EDSS 6, (E) EDSS 8, and (G) secondary progressive multiple sclerosis (SPMS). Forest plots for Cox proportional hazards models adjusted for relevant covariates for the time to (B) EDSS 4, (D) EDSS 6, (F) EDSS 8, and (H) SPMS.

measure for long term disability. rs10191329[A] was not associated with long-term disability (Fig. 1D). To increase statistical power, we next included the entire cohort of 1,455 pwMS. Once again, we did not detect any differences for last ARMSS score, age at onset between rs10191329 genotypes (Fig. 2A,B), or percentage of women (Fig. 2C) in the non-risk allele carrier group. Finally, we found that the risk genotype of rs10191329 was unable to predict ARMSS in the entire cohort (Fig. 2D).
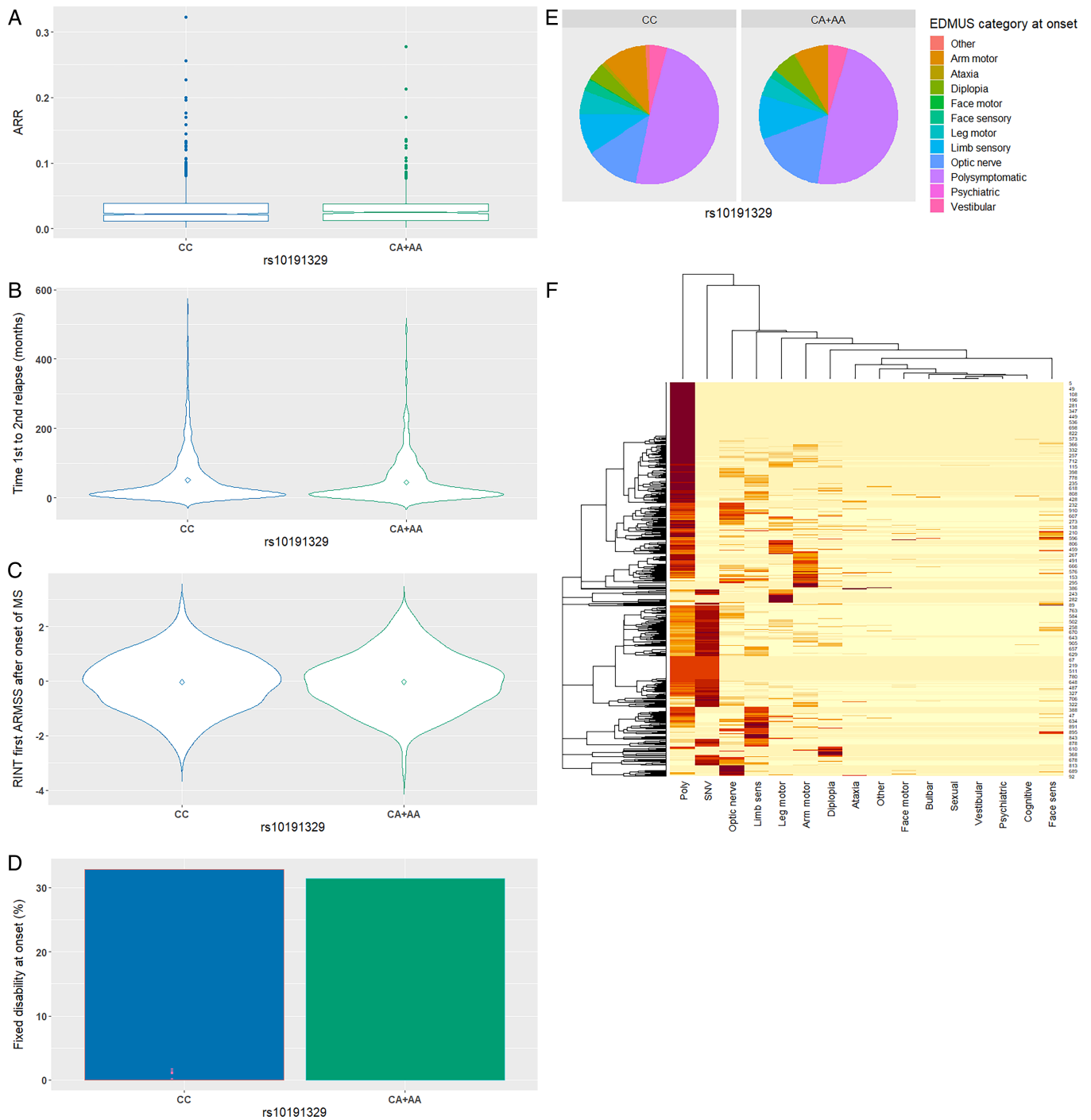
Figure 4: No significant differences in relevant disease outcomes in people with relapsing onset multiple sclerosis comparing risk carriers versus non-risk carriers for rs10191329. rs10191329 risk and non-risk carriers were compared for (A) annualized relapse rate, (B) time to second clinical relapse, (C) first age-related multiple sclerosis severity score (ARMSS) after onset of disease, (D) distribution of patients with and without fixed disability after their first relapse, (E) localization of onset of disease, and (F) distribution of the relapses according to anatomical localization.

Polygenic Risk Scores Not Associated With MS Severity

Although the most significant SNV from the IMSGC GWAS was not associated with severity in our dataset, it is possible that there may have been some genetic overlap for severity with studies for MS susceptibility loci. We, therefore, calculated a wGRS for MS susceptibility using minor allele counts in the lead SNVs at 181 genome-wide significant non-HLA loci, weighted by their effect size and an HLA genetic burden score (HLAGB) of 10 genome-wide significant SNV using the previous published odds ratio from the largest MS susceptibility GWAS.[2] Neither

the wGRS for MS susceptibility nor the HLAGB (Table **S1**) were significantly associated with ARMSS. These results indicate that MS susceptibility loci are not involved in the development of sustained disability. We then calculated a wGRS of 5 suggestive ($p < 5*10^{-6}$) MS severity SNVs, and weighted them according to the IMSGC severity GWAS.[6] The MS severity wGRS was not associated with the first ARMSS after disease onset (Fig. 2E), the last ARMSS (Fig. 2F), or the age at onset (Fig. 2G). No significant difference in the MS severity wGRS between relapsing onset and progressive onset patients was observed (Fig. 2H).

No Differences in Time to EDSS Milestones Between Different Genotypes of rs10191329

We assessed whether the MS severity wGRS was associated with time to EDSS milestones. We compared the pwMS with the highest and the lowest quantile of the MS severity wGRS, and found no effect of this PRS on time to EDSS 4 (maximum walking distance 500 meters without rest), EDSS 6 (requirement of unilateral walking aid to walk for 100 meters), EDSS 8 (restricted to wheelchair or bed),[15] or time to the development of secondary progressive MS in

pwMS with a relapsing onset[16] (all $P \geq 0.29$, Fig. S5). Next, we used survival analysis to assess whether the rs10191329 genotype is associated with speed to develop walking disability. Rs10191329 genotype was not associated with time to EDSS milestones or time to secondary progressive MS in pwMS with a relapsing onset, adjusted for sex, age at EDSS, and date of birth in our cohort of pwMS (Fig. 3A–G, all $P > 0.28$).[6] Survival analysis using an allelic approach showed similar non-significant findings (Fig. S3).

rs10191329[A] Not Associated With Clinical Phenotype of pwMS

We also investigated whether carrying the risk allele of rs10191329 was predictive of relapse activity in those with relapse-onset MS (n = 1,267). No differences were found between carriers and non-carriers of rs10191329[A] in ARR (Fig. 4A), time from first to second clinical episode of demyelination (Fig. 4B), level of disability after the first episode of demyelination (rank-inversed normalized ARMSS score, Fig. 4C), percentage of patients with incomplete recovery after their initial attack (Fig. 4D), anatomical localization of the first attack (Fig. 4E), or the distribution of
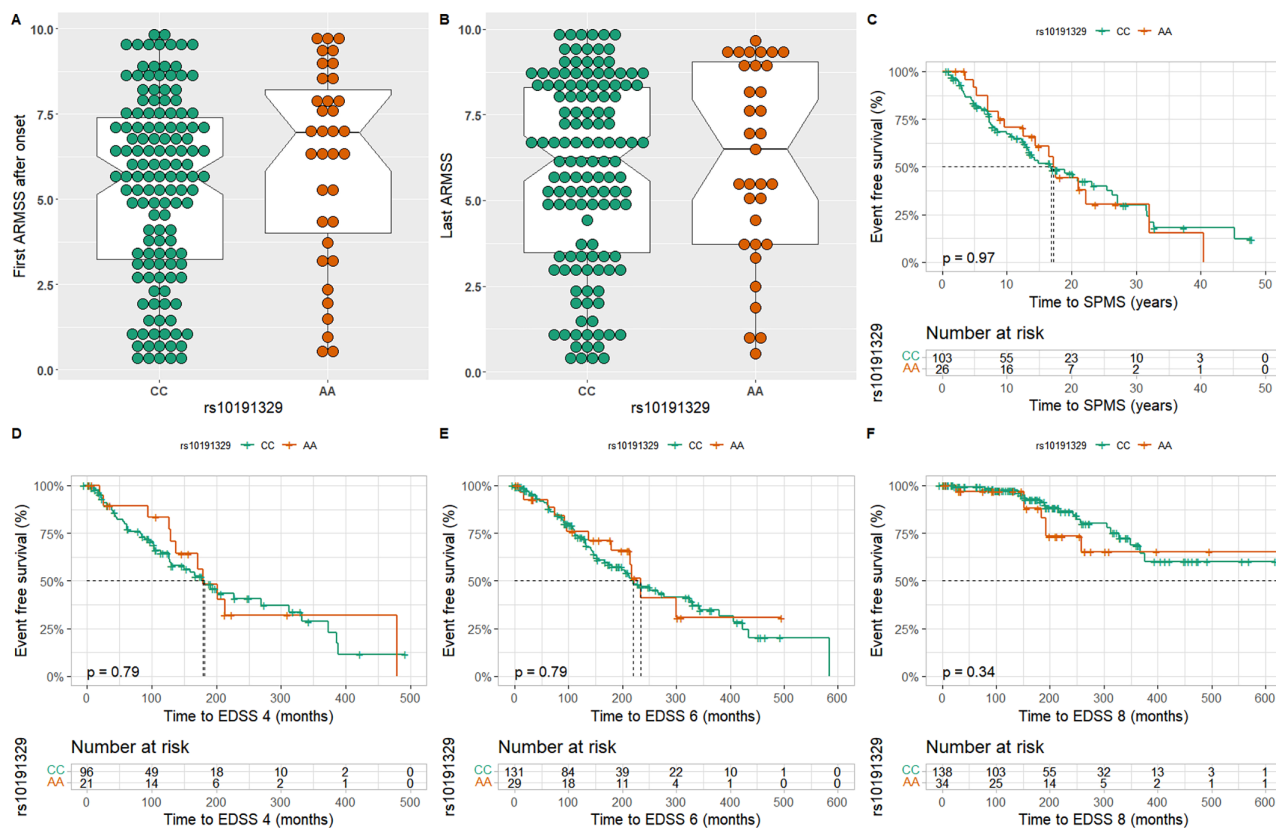
Figure 5: No differences in disease course and development of disability between homozygous risk carriers and non-risk carriers in a propensity matched case–control approach. Comparing rs10191329[CC] (homozygous non-risk carriers) with rs10191329[AA] (homozygous risk carriers) (A) first age-related multiple sclerosis severity score (ARMSS) after disease onset, (B) last known ARMSS, (C) time to secondary progressive multiple sclerosis (SPMS), (D) time to Expanded Disability Status Score (EDSS) 4, (E) time to EDSS 6, and (F) time to EDSS 8.

anatomical localizations of all experienced relapses of pwMS throughout the entire disease course (Fig. 4F).

Next, we applied an extremes of outcome approach to determine whether rs10191329[A] is associated with long-term disability development. We applied propensity score matching to 35 pwMS carrying rs10191329[AA] (2 rs10191329[AA] carriers were excluded due to missing data for propensity score matching) and 140 pwMS rs10191329[CC] matched for sex, age at onset, relapsing versus progressive onset, and use of disease-modifying drugs. To maximize statistical power, we applied a 1:4 case-to-control ratio.[14] No differences were found in the level of disability after disease onset (Fig. 5A), or the last known ARMSS score (Fig. 5B) between rs10191329[AA] and rs10191329[CC]. Next, we assessed whether rs10191329[AA] carriers more rapidly or frequently develop sustained disability compared with rs10191329[CC]. Time to EDSS 4, 6, and 8, and secondary progressive MS (only in relapsing onset pwMS) were similar between both genotypes (Fig. 5C–F, all P > 0.34). Finally, we assessed whether rs10191329[C] carriership is more common in benign MS. Therefore, we selected all pwMS with follow-up duration of at least 10 years, and compared the genotypes between patients who reached and who did not reach EDSS 4 using logistic regression analysis.

rs10191329[A] was not associated with a more aggressive disease course (OR 0.87, 95% CI 0.67–1.13, $p = 0.30$), whereas as expected, female sex was associated with a more favorable disease course (OR 0.64, 95% CI 0.48–0.84, $p = 0.002$) and a younger age at onset as well (OR 1.01, 95% CI 1.00–1.03, $p = 0.013$).

Two Previously Identified MS Severity SNVs Are Associated With Long-Term Outcomes in MS

Subsequently, we assessed whether we could validate the MSBase-identified variants associated with long-term disease outcomes.[7] We applied similar linear regression models, although DMT use in our model was classified as never exposed, only exposed to ME or only to HE, or switched during follow-up, rather than time on DMT used in the MSBase study. We validated rs7289446[G] as a SNV associated with longitudinal ARMSS scores. rs1207401 is in perfect linkage disequilibrium with rs7289446 ($R^2 = 1.0$, $D' = 1.0$) in the northern europeans from utah (CEU) population. Another suggestive SNV, rs868824[C], was also associated with MS severity score and a trend toward significance with ARMSS in our cohort (Table 2). The direction of the effect size of all SNVs was similar to the MSBase study. The

Table 2. Replication of suggestive multiple sclerosis Base severity single-nucleotide variant

	ARMSS				MSSS			
	Unadjusted		**Adjusted**		**Unadjusted**		**Adjusted**	
SNV	**Beta (SE)**	**p value**	**Beta (SE)**	**p value**	**Beta (SE)**	**p value**	**Beta (SE)**	**p value**
rs7289446								
AA	Ref							
GA	−0.10 (0.057)	0.08	−0.087 (0.056)	0.012	−0.08 (0.16)	0.62	−0.15 (0.15)	0.34
GG	**−0.27 (0.11)**	**0.01**	**−0.22 (0.10)**	**0.03**	−0.53 (0.29)	0.07	**−0.62 (0.28)**	**0.03**
rs1207401								
GG	Ref							
AG	−0.097 (0.057)	0.09	−0.087 (0.06)	0.12	−0.09 (0.16)	0.59	−0.15 (0.15)	0.32
AA	**−0.26 (0.11)**	**0.01**	**−0.21 (0.10)**	**0.04**	−0.50 (0.29)	0.09	**−0.58 (0.28)**	**0.04**
rs868824								
TT	Ref							
CT	0.024 (0.06)	0.70	0.042 (0.059)	0.48	0.19 (0.17)	0.26	0.17 (0.16)	0.30
CC	0.13 (0.08)	0.11	0.15 (0.08)	0.054	0.44 (0.23)	0.056	**0.44 (0.22)**	**0.04**

Note: The adjusted model included rs3135388 (proxy single-nucleotide variant for HLA-DRB1*1501), the first 5 principal components, the use of (never, only moderate efficacy, only high-efficacy or switching from moderate to high efficacy treatment), sex, and age at onset of multiple sclerosis as covariates. ARMSS = age-related multiple sclerosis severity score; MSSS = multiple sclerosis severity score; Ref = reference group SNV = single-nucleotide variant.

Table 3. HLA-DRB1*1501 carriership is associated with age at onset, but not with long-term disability outcomes

| | Heterozygous HLA-DRB1*1501 carrier | | | | Homozygous HLA-DRB1*1501 carrier | | | |
| | Unadjusted | | Adjusted | | Unadjusted | | Adjusted | |
Outcome	Beta (SE)	*p* value	Beta (SE)	*p* value	Beta (SE)	*p* value	Beta (SE)	*p* value
ARMSS	0.06 (0.06)	0.26	0.05 (0.05)	0.38	0.15 (0.11)	0.18	0.09 (0.11)	0.20
MSSS	−0.15 (0.16)	0.36	−0.02 (0.15)	0.87	−0.30 (0.31)	0.33	−0.03 (0.29)	0.93
Age at onset	−1.08 (0.59)	0.07	−1.07 (0.59)	0.07	−2.95 (1.16)	**0.01**	−2.70 (1.15)	**0.02**

Note: Results of linear regression to predict the outcome. Homozygous non-carriers of HLA-DRB1*1501 are the reference group. Age-related multiple sclerosis severity score and multiple sclerosis severity score were adjusted for the first five principal components, disease-modifying treatment usage as described, sex, and age at onset. The model to predict age at onset was adjusted for the same variables excluding age at onset as a covariate. Abbreviations: ARMSS = age-related multiple sclerosis severity score; DMT = disease-modifying treatment.

remaining suggestive MSBase severity SNVs were not significant in our cohort, and we could not replicate the sex-specific genetic effects (data not shown).

In addition, Brownlee et al demonstrated in a small cohort that HLA-DRB1*1501 carriership is associated with MS severity,[17] although the MSBase study was unable to replicate this finding.[7] We assessed the role of the tagging SNV rs3135388 (proxy for HLA-DRB1*1501) in our cohort. We were unable to identify an association between HLA-DRB1*1501 carriership and ARMSS or MS severity scores, although were able to replicate that HLA-DRB1*1501 carriership is associated with younger age at onset (Table 3).

Finally, we assessed whether there was an interaction between genotype and treatment modalities on the development of long-term fixed disability. We could not find an association between the SNV associated with disease progression and treatment responses (Table S3).

Discussion

Using a large MS cohort containing detailed prospective, longitudinal clinical data collected since 1985, we were unable to replicate the important finding of rs10191329[A] as a SNV associated with MS disease phenotype or severity. There could be several explanations for this lack of replication. First, all of our patients were participating in a prospective cohort study, rather than a case–control study. This limits the risk of selection bias; for example, all of the pwMS underwent EDSS measurement as part of routine clinical care. Therefore, the present cohort might be more representative of the general population of pwMS. However, in a subanalysis using a case–control approach based on genotype, we applied propensity matching to avoid confounding, and found similar non-significant findings.

Second, we only included EDSS scores measured during a physical consultation. The use of phone-based or interview-based EDSS outcomes is more reliable when patients have higher EDSS scores (≥4) compared with lower EDSS, which relies on a full neurological examination.[18] Non-physical EDSS scores may skew the inclusion of patients toward more severely affected patients, resulting in measurement bias. Finally, patients participating in the IMSGC study are selected based on EDSS criteria and repeated measurement, some of which may be derived from clinical trials, who are also at risk of selection bias. In the present study, all pwMS were derived from a longitudinal prospective cohort study, which is less prone to bias.

In the current study, rs10191329 was associated with an effect size of −0.06 (95% CI −0.16-0.05) for predicting ARMSS score (Fig. 2D). In other words, each copy of the A allele was associated with a reduction of 0.06 in ARMSS. Conversely, in the discovery phase of the IMSGC progression GWAS, an effect size of 0.089 was observed, and in the replication phase 0.044.[6] The lower effect size may be explained by a regression to the mean phenomenon,[19] and is likely to be a more accurate representation of the true effect size. Our effect size appears to be the opposite of the IMSGC study. However, a two-sided test for inference comparing both effect sizes was not statistically significantly different (*p* = 0.067), probably due to our smaller cohort size. The present study differs from the IMSGC study in several ways that might potentially result in differences in the observed genetic associations; our cohort likely includes a higher proportion of lower EDSS scores due to our prospective study design. Additionally, the density of EDSS measures is higher in more recent cases compared with historical cohorts (due to increased monitoring on treatments) and, therefore, potentially shortening the time to EDSS 4.0. Finally, early

commencement on DMTs may improve disease trajectory. Alternatively, it could be that rs10191329 is dominantly involved in neurodegeneration rather than earlier inflammatory phases of disease, which are more commonly associated with lower EDSS scores.

Recently, MSBase published their progression GWAS and found rs7289446 was associated with ARMSS, and rs1207401 with MS severity scores.[7] We replicated these SNV associated with sustained disability in MS. In both the MSBase and the present study, a wGRS of MS susceptibility genes was not associated with disease severity. In another small cohort, HLA-DRB1*1501 was associated with disability in MS,[17] whereas in the MSBase study, HLA-DRB1*1501 was only associated with age at onset, similar to findings in the present study.

Measuring long-term disability in MS is complex and difficult, and historically has tended to rely mainly on EDSS measurement. This nonlinear scale is heavily dependent on assessment of ambulation in the higher scores, and is less sensitive in capturing other aspects of disability, such as reduced dexterity, cognition, or hand function.[20] It is also well-known that relapses can impact the development of disability, mainly by a temporary worsening of neurological function, rather than a long-term effect on the development of sustained disability.[21] We compared carriers of rs10191329[A] with rs10191329[C] and found no differences in the ARR (Fig. 4A). In addition, we did not observe any difference in anatomical localization of the relapses between carriers and non-carriers of rs10191329 (Fig. 4E,F). Finally, the percentage of patients who did not completely recover after the first onset of demyelination was similar between the genotypes (Fig. 4D). Therefore, it is unlikely that rs10191329 affects the development of sustained disability via relapses. Earlier studies assessing MS disability failed to find genetic associations, partially due to the relatively low number of pwMS included.[22,23] Several studies found that HLA-DRB1*1501 is associated with a lower age at onset, and that age at onset on itself is associated with the development of long-term disability.[24,25] Of interest, in the present study, HLA-DRB1*1501 carriership was significantly higher in rs10191329[A] carriers (Table 1) and, therefore, it would be of interest to assess in the IMSGC study whether HLA-DRB1*1501 may affect age at onset and, thereby, the development of disease severity.

Several further limitations of the present study should be acknowledged. First, the number of homozygous risk carriers of rs10191329 is relatively low. We applied several different analysis strategies on multiple important clinical outcomes associated with disease severity and phenotype, including an extremes of outcome analysis after propensity score matching on homozygous risk and non-risk carriers with the optimal case-to-control ratio of 1:4 to maximize statistical power.[14] All analyses showed similar non-significant results without a trend toward significance, limiting the application of rs10191329 in clinical prognostication of pwMS. The current study was a candidate gene study, and, therefore, had no requirement to correct for multiple testing. Second, after imputation and quality control, the suggestive SNV rs149097173 of the IMSGC study was not available for further study in our cohort. Finally, the EDSS data were obtained over a time period from 1985 to 2022; therefore, several different experienced and trained examiners determined those scores. It is well-known that the interrater variability of EDSS is relatively high, and we could not exclude that this may slightly impact the results,[26] although this would also reflect the approach by the IMSGC study.

In conclusion, the severity GWAS by the IMSGC is the largest study showing two genetic loci associated with MS severity, whereas the smaller MSBase GWAS found 10 suggestive SNVs, which did not reach genome-wide significance. The lack of association with disease phenotype in the present cohort for the IMSGC GWAS and the replication of two of the suggestive MSBase GWAS hits shows the importance of validation in independent cohorts to further understand the genetic basis of neurodegeneration. Replication of variants associated with disease severity in neurodegenerative disorders with complex phenotypes remains challenging,[27,28] and standardization of data and methodology will aid future analysis. In the meantime and considering the relatively small effect sizes of the MS severity SNVs reported to date, it seems unlikely that these variants will be informative in management decisions or patient counseling.

Author contributions

K.L.K., E.C.T., and NPR contributed to the conception and design of the study. K.L.K., E.U., S.L., M.W., R.W.T., K.E.H., E.C.T., and N.P.R. contributed to the acquisition and analysis of data. K.L.K., E.U., S.L., M.W., R.W.T., K.E.H., P.H., M.L., E.C.T., and N.P.R. contributed to drafting the text or preparing the figures.

Acknowledgment

Potential Conflicts of Interest

Nothing to report.

References

1. Compston A, Coles A. Multiple sclerosis. Lancet 2002;359:1221–1231.

2. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science 2019;365:eaav7188.

3. George MF, Briggs FBS, Shao X, et al. Multiple sclerosis risk loci and disease severity in 7,125 individuals from 10 studies. Neurol Genet 2016;2:e87.

4. International Multiple Sclerosis Genetics Consortium. Genome-wide association study of severity in multiple sclerosis. Genes Immun 2011;8:615–625.

5. Mowry EM, Carey RF, Blasco MR, et al. Multiple sclerosis susceptibility genes: associations with relapse severity and recovery. PloS One 2013;8:e75416.

6. International Multiple Sclerosis Genetics Consortium, MultipleMS Consortium. Locus for severity implicates CNS resilience in progression of multiple sclerosis. Nature 2023;619:323–331.

7. Jokubaitis VG, Campagna MP, Ibrahim O, et al. Not all roads lead to the immune system: the genetic basis of multiple sclerosis severity. Brain 2023;146:2316–2331.

8. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 2018;17:162–173.

9. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7.

10. Manouchehrinia A, Westerlind H, Kingwell E, et al. Age related multiple sclerosis severity score: disability ranked by age. Mult Scler 2017;23:1938–1946.

11. Filippi M, Amato MP, Centonze D, et al. Early use of high-efficacy disease-modifying therapies makes the difference in people with multiple sclerosis: an expert opinion. J Neurol 2022;269:5382–5394.

12. Confavreux C, Compston DA, Hommes OR, et al. EDMUS, a European database for multiple sclerosis. J Neurol Neurosurg Psychiatry 1992;55:671–676.

13. De Jager PL, Chibnik LB, Cui J, et al. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurol 2009;8:1111–1119.

14. Hennessy S, Bilker WB, Berlin JA, Strom BL. Factors influencing the optimal control-to-case ratio in matched case-control studies. Am J Epidemiol 1999;149:195–197.

15. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33:1444–1452.

16. Lorscheider J, Buzzard K, Jokubaitis V, et al. Defining secondary progressive multiple sclerosis. Brain 2016;139:2395–2405.

17. Brownlee WJ, Tur C, Manole A, et al. *HLA-DRB1*1501* influences long-term disability progression and tissue damage on MRI in relapse-onset multiple sclerosis. Mult Scler 2023;29:333–342.

18. Bove R, Bevan C, Crabtree E, et al. Toward a low-cost, in-home, telemedicine-enabled assessment of disability in multiple sclerosis. Mult Scler 2019;25:1526–1534.

19. Stigler SM. Regression towards the mean, historically considered. Stat Methods Med Res 1997;6:103–114.

20. Weinstock-Guttman B, Sormani MP, Repovic P. Predicting long-term disability in multiple sclerosis: a narrative review of current evidence and future directions. Int J MS Care 2022;24:184–188.

21. University of California, San Francisco MS-EPIC Team, Cree BAC, Hollenbach JA, et al. Silent progression in disease activity-free relapsing multiple sclerosis. Ann Neurol 2019;85:653–666.

22. Baranzini SE, Wang J, Gibson RA, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet 2009;18:767–778.

23. Jensen CJ, Stankovich J, Van der Walt A, et al. Multiple sclerosis susceptibility-associated SNPs do not influence disease severity measures in a cohort of Australian MS patients. PloS One 2010;5:e10003.

24. Isobe N, Keshavan A, Gourraud PA, et al. Association of HLA genetic risk burden with disease phenotypes in multiple sclerosis. JAMA Neurol 2016;73:795–802.

25. Masterman T, Ligers A, Olsson T, et al. HLA-DR15 is associated with lower age at onset in multiple sclerosis. Ann Neurol 2000;48:211–219.

26. Cohen M, Bresch S, Thommel Rocchi O, et al. Should we still only rely on EDSS to evaluate disability in multiple sclerosis patients? A study of inter and intra rater reliability. Mult Scler Relat Disord 2021;54:103144.

27. Iwaki H, Blauwendraat C, Leonard HL, et al. Genomewide association study of Parkinson's disease clinical biomarkers in 12 longitudinal patients' cohorts. Mov Disord 2019;34:1839–1850.

28. Martínez Carrasco A, Real R, Lawton M, et al. Genome-wide analysis of motor progression in Parkinson disease. Neurol Genet. 2023;9:e200092.