Received: 26 April 2021 | Accepted: 9 July 2021

DOI: 10.1111/ene.15019

european journal of neurology
the official journal of the european academy of neurology

ORIGINAL ARTICLE

T cell composition and polygenic multiple sclerosis risk: A population-based study in children

Casper L. de Mol[1,2] | Kirsten I. M. Looman[2,3] | Marvin M. van Luijn[4] | Karim L. Kreft[1] | Philip R. Jansen[2,5,6] | Menno C. van Zelm[7] | Joost J. F. M. Smolders[1,4] | Tonya J. H. White[8,9] | Henriette A. Moll[2,3] | Rinze F. Neuteboom[1]

[1]Department of Neurology, MS Center ErasMS, Erasmus University Medical Center, Rotterdam, the Netherlands

[2]Generation R Study Group, Erasmus University Medical Center, Rotterdam, the Netherlands

[3]Department of Pediatrics, Erasmus University Medical Center, Rotterdam, the Netherlands

[4]Department of Immunology, Erasmus University Medical Center, Rotterdam, the Netherlands

[5]Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Amsterdam University Medical Center, Amsterdam, the Netherlands

[6]Department of Clinical Genetics, Amsterdam University Medical Center, Amsterdam, the Netherlands

[7]Department of Immunology and Pathology, Central Clinical School, Monash University and Alfred Hospital, Melbourne, Victoria, Australia

[8]Department of Child and Adolescent Psychiatry, Erasmus University Medical Center, Rotterdam, the Netherlands

[9]Department of Radiology and Nuclear Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands

Correspondence
Rinze F. Neuteboom, Erasmus MC-Sophia, Room SK-1210, PO Box 2060, 3015 GD Rotterdam, the Netherlands.
Email: r.neuteboom@erasmusmc.nl

Funding information
This study was sponsored by the Dutch MS Research Foundation. The Generation R study is made possible by continuous support from the Erasmus University Medical Center, Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development, the Netherlands Organization for Scientific Research, and the Ministry of Health, Welfare, and Sport.

Abstract

Background and purpose: Patients with multiple sclerosis (MS) have altered T cell function and composition. Common genetic risk variants for MS affect proteins that function in the immune system. It is currently unclear to what extent T cell composition is affected by genetic risk factors for MS, and how this may precede a possible disease onset. Here, we aim to assess whether an MS polygenic risk score (PRS) is associated with an altered T cell composition in a large cohort of children from the general population.

Methods: We included genotyped participants from the population-based Generation R study in whom immunophenotyping of blood T cells was performed at the age of 6 years. Analyses of variance were used to determine the impact of MS-PRSs on total T cell numbers ($n = 1261$), $CD4^+$ and $CD8^+$ lineages, and subsets therein ($n = 675$). In addition, T-cell-specific PRSs were constructed based on functional pathway data.

Results: The MS-PRS negatively correlated with $CD8^+$ T cell frequencies ($p = 2.92 \times 10^{-3}$), which resulted in a positive association with $CD4^+/CD8^+$ T cell ratios ($p = 8.27 \times 10^{-9}$). These associations were mainly driven by two of 195 genome-wide significant MS risk variants: the main genetic risk variant for MS, *HLA-DRB1*15:01* and an *HLA-B* risk variant. We observed no significant associations for the T-cell-specific PRSs.

Conclusions: Our results suggest that MS-associated genetic variants affect T cell composition during childhood in the general population.

Casper de Mol is sole first author on the paper. Kirsten I. M. Looman and Marvin M. van Luijn are shared second author on the paper and have contributed equally.

KEYWORDS
Epstein–Barr virus infections, genetic association studies, multiple sclerosis, T lymphocytes, vitamin D

INTRODUCTION

The exact pathophysiology of the autoimmune response in multiple sclerosis (MS) remains to be elucidated, and results from a complex interplay between genetic and environmental risk factors that contribute to disease risk. Migration studies suggest a pivotal time window during childhood in which environmental risk factors (e.g., Epstein–Barr virus [EBV] or cytomegalovirus [CMV] seropositivity, and serum 25-hydroxyvitamin-D [25(OH)D]) contribute to MS risk [1,2].

Genome-wide association studies (GWASs) have identified 233 genetic variants (single nucleotide polymorphisms [SNPs]) that significantly affect disease susceptibility in adult patients [3]. These studies have shown that the genetic risk of MS is polygenic, with a large number of genetic variants that each have a small effect on disease risk [3]. However, the majority of people with high genetic risk for MS are never diagnosed with the disease [4]. How these risk variants together affect biological mechanisms culminating in a higher risk of developing MS and at which time points during the lifespan are unknown. To capture this combined effect of genetic risk variants, polygenic risk score (PRS) analyses can be used [5].

T cells are suggested to be the main immune lineage that are directly or indirectly influenced by genetic MS risk variants [3]. Several studies have described altered T cell homeostasis in MS patients, with increased $CD4^+/CD8^+$ ratios that decrease during treatment [6,7].

Here, we aimed to determine how PRSs for MS are associated with the distribution of naive, memory, and activated $CD4^+$ and $CD8^+$ T cell populations in 6-year-old children from the general population. We assessed the main genetic variants driving these associations and the relationship between overall genetic MS risk and EBV and CMV seropositivity, as well as serum 25(OH)D levels.

MATERIALS AND METHODS

Study sample

This study was performed in the context of the population-based Generation R study (n = 9749), located in the Netherlands, which investigates various aspects of childhood development [8]. At the age of 6 years, absolute numbers of total T cells were measured in 3465 participants, and detailed phenotyping of T cell subsets was performed in 1079 participants [9]. In the current study, we included unrelated participants of European ancestry, who had high-quality genotype data available (Figure 1).

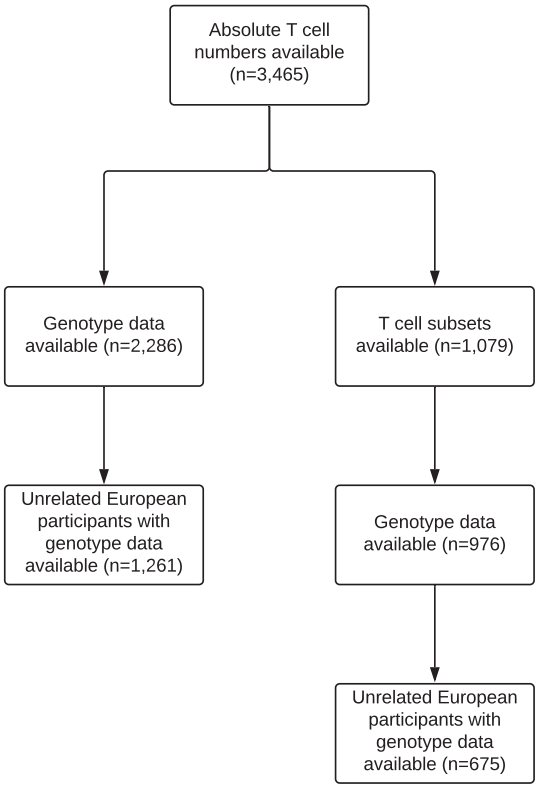

FIGURE 1 Flowchart of the participant selection in the current study

The study protocol of the Generation R study has been approved by the Medical Ethical Committee of the Erasmus Medical Center in Rotterdam. The legal representatives of the children provided written informed consent.

Genotype data

Sample collection of the genetic data and genotype calling procedures have been reported in previous work [10]. Genotype data were collected at birth, derived from cord blood, or during a visit to the research center, and genotyped on Illumina 610K and 660K SNP arrays. Information on the quality control of the genetic data, the imputation method, and the calculation of principal components can be found elsewhere [11]. To summarize, we selected subjects of European ancestry based upon the first four principal components inside the range of the HapMap Phase II Northwestern European founder population [12]. Furthermore, we imputed the genotype data using 1000 Genomes (Phase I, Version 3) data to calculate our PRSs [13].

Polygenic scoring

We used a large discovery GWAS for MS (N = 41,505; 14,802 cases/26,703 controls), carried out by the International Multiple Sclerosis Genetics Consortium (IMSGC; https://imsgc.net/), using imputed genotype data to estimate weighted PRSs [3]. We used PRSice 2 [14], an R-script to calculate clumped (r^2 < 0.10, window = 250 kB) and p-value thresholded polygenic scores in PLINK (v1.9) [15], for the computation of our PRSs. We used various p-value thresholds (P_T < 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1) in the calculation of the PRSs to include the effect of suggestive variants and capture the polygenic architecture of MS [3]. In addition, we computed a MS-PRS only incorporating the genome-wide risk variants ($P_T < 5 \times 10^{-8}$) from the final meta-analysis of the IMSGC (47,429 cases/68,374 controls).

Because of the strong effect of the major histocompatibility complex (MHC) region on MS pathophysiology [3], three sets of PRSs were calculated: one with all SNPs included, one excluding SNPs from the MHC region, and one only including SNPs from the MHC. To characterize the presence of *HLA-DRB1*15:01*, we used rs3135388 as tag SNP [16]. Rs9266629 was included in our analyses as tag SNP for *HLA-B* genetic risk variant rs3819284. In addition, we used rs9268839, rs1057149, and rs2187688 as tag SNPs for *HLA-DRB* variants of importance in rheumatoid arthritis, diabetes mellitus type 1, and systemic lupus erythematosus to investigate possible confounding of the T cell distribution [17,18].

Due to the low minor allele frequency (<0.20) of rs3135388 and rs9266629, we dichotomized their carriership into two groups, homozygous non-risk and at-risk participants. At-risk participants consisted of heterozygous and homozygous carriers of the effect allele.

Construction of T-cell-specific PRSs

In a final analysis, we used cis expression quantitative trait loci (cis-eQTL) effect data to associate the MS genome-wide risk variants with particular T cell subsets, providing us with the possibility to construct PRSs specific for these subsets.

$CD4^+$ and $CD8^+$ T-cell-specific PRSs were constructed by assessing whether a locus implicated by an SNP [3] had high mRNA expression in a specific subset [19–22]. Additionally, eQTL were assessed for all MS risk SNPs and assigned to specific T cell subsets [20]. Lastly, gene ontology analyses were performed on all MS risk SNPs to assign loci to functional pathways in the aforementioned lymphocyte subsets [23–25].

EBV and CMV serology

To identify the presence of immunoglobulin G antibodies against CMV and EBV-viral capsid antigen (VCA), enzyme-linked immunoassays were used, as described before [9]. Seropositivity was defined by a sample-threshold ratio greater than 0.6 (CMV) and 0.8 (EBV capsid antigen).

Vitamin D assessment

Measurements of 25(OH)D were conducted in blood sera of children at the age of 6 years (110 µl per sample) using the liquid chromatography/tandem mass spectrometry method, which is a highly sensitive method for measuring 25(OH)D levels, commonly used in epidemiological studies [26]. Serum 25(OH)D levels were residualized for season of blood draw to adjust for the nonlinear effect of season on serum 25(OH)D levels.

T cell phenotyping

The presence and phenotype of $CD4^+$ and $CD8^+$ T cell populations were analyzed in whole blood using six-color flow cytometry [9]. Absolute T cell numbers were determined with a routine diagnostic lyse-no-wash protocol [9]. We analyzed the following subsets within the $CD4^+$ and $CD8^+$ T cell lineage [9]: naive ($CD45RO^-CCR7^+$), central memory (Tcm; $CD45RO^+CCR7^+$), CD45RO-positive effector memory ($CD45RO^+CCR7^-$), CD45RA-positive effector memory (TemRA; $CD45RO^-CCR7^-$), terminally differentiated ($CD57^+$), and activated ($HLA\text{-}DR^+$) T cells. $CD56^+$ natural killer T cells were excluded from our analysis. The above subsets were analyzed as part of the Generation R study [6,7]. $CD4^+/CD8^+$ T cell ratios were calculated by dividing the total number of $CD4^+$ and $CD8^+$ T cells. Flow cytometric data were obtained using an LSRII flow cytometer (BD Biosciences) with standardized instrument settings and analyzed using FACSDiva analysis software [27,28].

Statistical analyses

We performed our analyses using R statistical software (v3.5.1) [29]. Before our analyses, we divided the MS-PRSs into quartiles. In the first analyses, we analyzed the association between environmental risk factors for MS (EBV, CMV, and vitamin D) and our PRSs, to investigate possible confounding in our subsequent T cell analyses. To determine the association between the MS-PRSs' quartiles and EBV and CMV seropositivity, logistic regression was performed across the different p-value thresholds. We used analyses of variance (ANOVAs) to investigate the relationship between the MS-PRSs' quartiles and serological 25(OH)D levels.

All T cell outcomes of interest were log-transformed to ensure normally distributed residuals of our ANOVA models in the subsequent T cell analyses. At first, the MS-PRSs' quartiles, based on different thresholds, were all associated with absolute T cell counts using ANOVAs. Next, the PRS threshold with the strongest association (lowest p-value) was used in subsequent ANOVAs investigating the associations with different T cell subpopulations of interest. In

addition, we performed ANOVAs with the MS-PRS including only the genome-wide significant risk variants, to compare the effects of these variants on T cell populations with the effects of suggestive variants. Tukey tests were used to investigate post hoc differences in significant ANOVA associations.

In all our analyses, we adjusted for age at blood withdrawal, sex, and the first 10 genetic principal components. Effect modification by sex was tested in a separate sensitivity analysis. We used false discovery rate (FDR) per research question to account for multiple testing [30].

RESULTS

Participant characteristics

Of the participants for whom absolute T cell counts were determined, 2286 had genotype data available. After selection on relatedness, genotype quality, and European ancestry, 1261 participants remained eligible for analyses using absolute T cell numbers (Figure 1). Participants had an even distribution of sex (50.8% male) and a median age of 6.0 years (interquartile range = 5.9–6.2; Table 1). No difference was found in mean PRS for MS compared to the participants who had no absolute T cell counts available ($n = 1569$; $p = 0.62$).

TABLE 1 Demographics and laboratory measurements in study participants

Characteristic	Absolute T cell counts, $n = 1261$	T cell phenotyping, $n = 675$
Age, years, median (IQR)	6.0 (5.9–6.2)	6.0 (5.9–6.2)
Male, *n* (%)	641 (50.8)	340 (50.4)
Level of maternal education, *n* (%)		
High	818 (64.8)	419 (62.1)
Middle	405 (32.1)	239 (35.4)
Low	17 (1.3)	9 (1.3)
Unknown	21 (1.7)	8 (1.2)
Season of blood draw, *n* (%)		
Spring	282 (22.4)	171 (25.3)
Summer	290 (23.0)	150 (22.2)
Autumn	351 (27.8)	174 (25.8)
Winter	338 (26.8)	180 (26.7)
EBV seropositivity, *n* (%)	512/1225 (41.8)	270/649 (41.6)
CMV seropositivity, *n* (%)	353/1227 (28.8)	191/649 (29.4)
Serum 25(OH)D levels, nmol/L, median (IQR)	74.0 (58.7–88.0)	74.0 (57.0–88.0)
Reported presence of paternal MS, *n* (%)	2/1036 (0.2)	0/572 (0.0)
Reported presence of maternal MS, *n* (%)	2/1144 (0.2)	0/611 (0.0)

Abbreviations: 25(OH)D, 25-hydroxyvitamin-D; CMV, cytomegalovirus; EBV, Epstein–Barr virus; IQR, interquartile range; MS, multiple sclerosis.

Detailed immunophenotyping data were available from 1079 participants. Selection on relatedness, genotype quality, and ancestry left us with 675 participants eligible for the detailed T cell subset analyses (Figure 1). Descriptive characteristics of these participants were comparable to the participants with absolute T cell counts available (Table 1), and the MS-PRS did not differ from participants who did not take part in the detailed immunophenotyping ($n = 2155$; $p = 0.42$).

EBV and CMV serology and serum 25(OH)D levels

EBV and CMV serology was measured in 4464 participants. Of these participants, 1551 had good-quality genotype data available. We found no relationship in these participants between the PRSs' quartiles for MS and EBV and CMV seropositivity (Tables S1 and S2).

Serum 25(OH)D levels were measured in 3983 participants, of whom 1442 had good-quality genotype data available. The known association of the rs7041 A-allele with lower 25(OH)D levels was replicated in our dataset ($p = 5.19 \times 10^{-10}$) [31]. The genotype data did not include rs5688. The MS risk SNP located in the *CYP24A1* gene (rs2248137) and the MS-PRSs' quartiles across all thresholds were not associated with the level of serum 25(OH)D in these participants (Table S3).

Effects on absolute T cell counts

We found no significant association between the MS-PRSs' quartiles and absolute T cell counts across all thresholds (Table S4). The PRS with a threshold of $P_T < 0.005$ revealed the strongest association with absolute total T cell counts ($p = 0.29$) and was used in our subsequent detailed immunophenotyping analyses.

Associations with T cell subpopulations

The MS-PRS ($P_T < 0.005$) had a significant association with total ($p = 2.92 \times 10^{-3}$) and naive ($p = 1.85 \times 10^{-4}$) $CD8^+$ T cell numbers and $CD4^+/CD8^+$ T cell ratios ($p = 8.27 \times 10^{-9}$). A complete overview of the associations between the MS-PRS ($P_T < 0.005$) and T cell populations can be found in Table 2. Post hoc Tukey tests revealed significant negative associations of the first and second MS-PRSs' quartiles with the fourth quartile considering both total and naive $CD8^+$ T cell numbers (Figure 2a,b). Accordingly, this resulted in a significant positive association of the PRS with the $CD4^+/CD8^+$ ratio across multiple quartiles (Figure 2c). In the sensitivity analyses, we found no evidence for interaction effects by sex (Table S5).

To investigate the MHC dependency of these significant associations, we excluded the *HLA-DRB1*15:01* tag variant rs3135388 and the MHC region from our MS-PRS (Tables S6–S9). The removal of rs3135388 (and variants with high linkage disequilibrium to this variant within a 1-Mb region) attenuated the significant $CD8^+$

TABLE 2 Analysis of variance results of the multiple sclerosis polygenic risk score ($P_T < 0.005$) and T cell subsets of interest

Cell subset	Unadjusted *p*	FDR-adjusted *p*
Total $CD4^+$ T cells	0.17	0.71
$CD4^+$ Naive	0.14	0.70
$CD4^+$ Tcm	0.54	0.77
$CD4^+$ TemRO	0.49	0.77
$CD4^+$ TemRA	0.46	0.77
Total $CD8^+$ T cells	2.92×10^{-3}	0.02[a,*]
$CD8^+$ naive	1.85×10^{-4}	3.14×10^{-3}[a,*]
$CD8^+$ Tcm	0.25	0.72
$CD8^+$ TemRO	0.32	0.72
$CD8^+$ TemRA	0.61	0.77
$CD3^+CD8^-$ HLA-DR^+	0.53	0.77
$CD3^+CD8^+$ HLA-DR^+	0.27	0.72
$CD3^+CD8^+CD57^-CD56^-$	0.55	0.77
$CD3^+CD8^+CD57^+CD56^-$	0.56	0.77
$CD3^+CD8^-CD57^-CD56^-$	0.59	0.77
$CD3^+CD8^-CD57^+CD56^-$	0.83	0.91
$CD4^+/CD8^+$ ratio	8.27×10^{-9}	2.81×10^{-7}[a,**]

Note:: Included: $n = 675$ children; data are corrected for age, sex, and 10 genetic principal components.

Abbreviations: FDR, false discovery rate; Tcm, central memory T cells; TemRA, RA-positive effector memory T cells; TemRO, RO-positive effector memory T cells.

[a]Significant values after FDR multiple testing correction.

*$p < 0.05$, **$p < 0.001$.

results, but these associations remained significant and the association involving the $CD4^+/CD8^+$ ratio increased in significance. Additional adjustment for *HLA-DRB1*15:01* status within this model still resulted in a significant association between the MS-PRS and the $CD4^+/CD8^+$ ratio (Table S7). When adding the status of several other *HLA-DRB* alleles to our model as covariates, we again observed a significant association of the MS-PRS with the $CD4^+/CD8^+$ ratio (Table S8). After removal of the total MHC region, the association between the MS-PRS and $CD4^+/CD8^+$ ratio became weaker, but remained significant (Table S9). The PRS with only MHC risk variants showed comparable results to our PRS including both MHC and non-MHC risk variants (Table S10). Due to the different linkage disequilibrium structure of the MHC region, we constructed several PRSs using various clumping parameters for the MHC region. However, altered clumping of the MHC did not significantly change our results (Table S11).

In addition, we calculated an MS-PRS only incorporating risk variants that were genome-wide significant ($P_T < 5 \times 10^{-8}$) in the MS GWAS performed by the IMSGC [3], and analyzed the association with the T cell subsets (Table S12). We observed a similar positive association of the first and second quartiles of the MS-PRS with the fourth quartile concerning the $CD4^+/CD8^+$ ratio (Figure 2d and Table S12).

Individual risk variant analyses

Subsequently, we investigated the relationship between an increased $CD4^+/CD8^+$ ratio and individual genetic MS risk variants. Of the 195 of 233 MS genome-wide significant risk SNPs that were available in our study population, two risk variants showed a significant association with the $CD4^+/CD8^+$ ratio after FDR correction (Table S13, Figure 3a). A positive association was found for rs3135388 (*HLA-DRA*, a tag variant for *HLA-DRB1*15:01*), whereas a negative association was observed for rs9266629 (intergenic, between *ZDHHC20P2* and *FGFR3P1*, a tag variant for *HLA-B* variant rs3819284).

Next, we associated these two risk variants with the previous T cell subsets of interest (Tables S14 and S15). Rs3135388 status (MHC Class II) was positively associated with $CD4^+$ total and naive numbers, whereas rs9266629 (MHC Class I) showed positive associations with various $CD8^+$ subsets (e.g., total $CD8^+$, $CD8^+$ naive, and $CD8^+$ Tcm; Figure 3a). Figure 3b shows the combined effects of the two risk variants on the $CD4^+/CD8^+$ ratio.

T cell PRS

In a final analysis, the autosomal genome-wide significant MS risk variants from the IMSGC were classified into specific MS-associated T cell PRSs (Figure 4, Table S16). We observed no significant associations between the different PRSs and the T cell outcomes, including the $CD4^+/CD8^+$ ratio.

When incorporating rs3135388 and rs9266629 status into the T cell PRSs, we observed several dose-dependent significant associations (Tables S17–S19). A higher $CD8^+$ PRS quartile was associated with a lower count of $CD8^+$ T cell subsets and a higher $CD4^+/CD8^+$ ratio ($p = 8.55 \times 10^{-7}$). The PRS involving Th1-, Th17-, Th17.1-, and $CD8^+$-associated variants showed positive associations with $CD4^+$ TemRA counts ($p = 0.01$) and the $CD4^+/CD8^+$ ratio ($p = 2.34 \times 10^{-6}$; Table S18). Similar associations were observed for the PRS incorporating variants associated with Treg and Th2 cells (Table S19). All these associations were driven by rs3135388 and rs9266629 status.

DISCUSSION

In this study, we found that PRSs for MS are associated with T cell numbers in the peripheral blood at the age of 6 years in children from the general population. A higher genetic risk for MS was associated with an increased $CD4^+/CD8^+$ ratio, which was the result of a negative association between the MS-PRS and total and naive $CD8^+$ T cell numbers. Two genome-wide significant MS risk variants, rs3135388 and rs9266629, were individually associated with the changes in the $CD4^+/CD8^+$ ratio and absolute numbers of several $CD4^+$ and $CD8^+$ T cell subsets. No associations were found between T-cell-specific MS-PRSs and T cell numbers.

We found no associations between MS-PRSs and seropositivity for EBV and CMV, and serum 25(OH)D levels. We conclude that

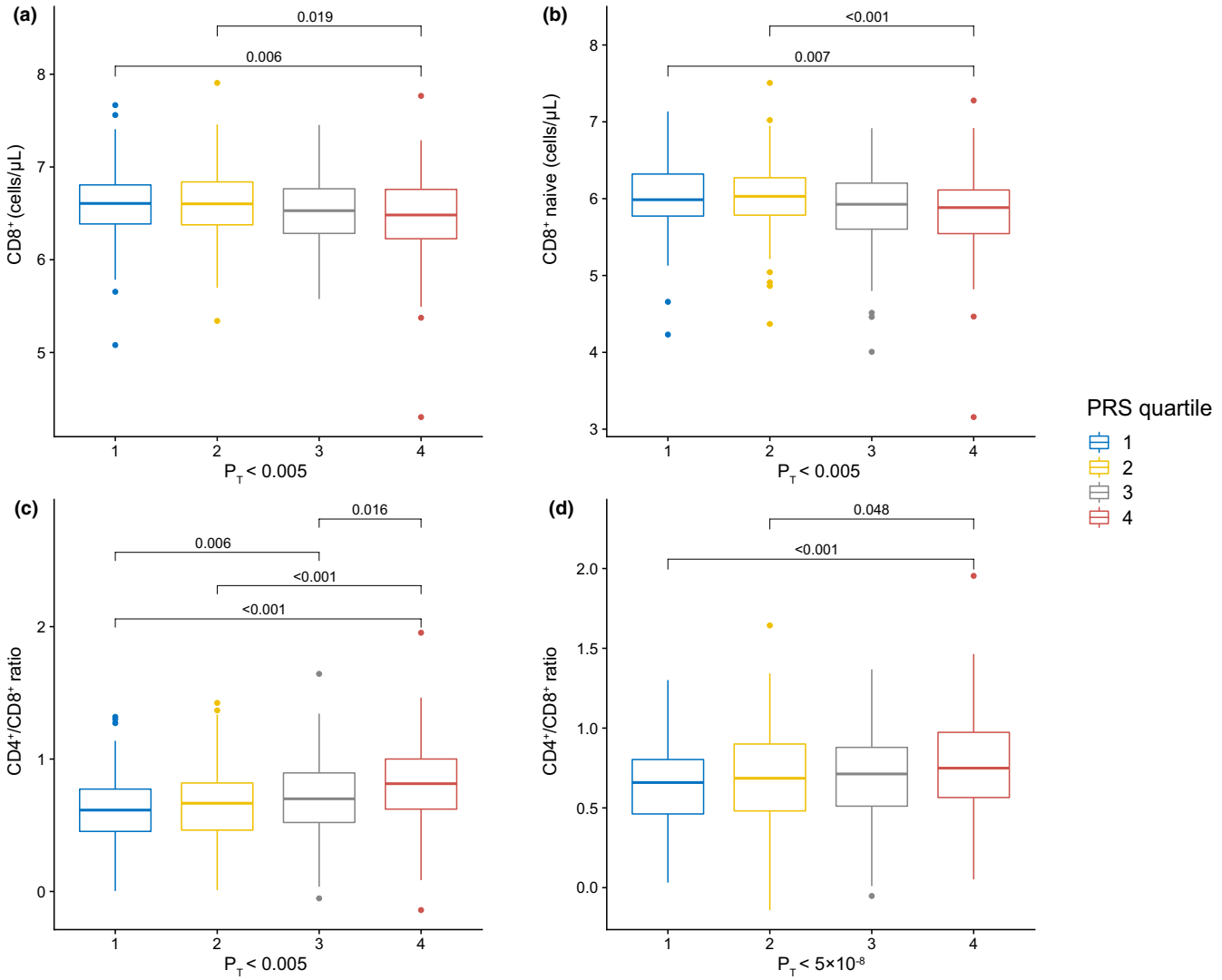

FIGURE 2 Post hoc Tukey test results of the different polygenic risk scores (PRSs) associated with the $CD4^+/CD8^+$ ratio and other T cell subsets. Numbers on the x-axis correspond to the different quartiles of the multiple sclerosis (MS)-PRS. Based on increasing MS-PRS, the participants were divided into four quartiles of similar sample size. (a) Association between total $CD8^+$ cells and the quartiles of the MS-PRS ($P_T < 0.005$). (b) Association between naive $CD8^+$ cells and the quartiles of the MS-PRS ($P_T < 0.005$). (c) Association between the $CD4^+/CD8^+$ ratio and the quartiles of the MS-PRS ($P_T < 0.005$). (d) Association between the $CD4^+/CD8^+$ ratio and the quartiles of the MS-PRS ($P_T < 5 \times 10^{-8}$) [Colour figure can be viewed at wileyonlinelibrary.com]

MS-PRSs are associated with T cell composition in children of a general population, and that this mechanism likely contributes to the overall risk of developing MS later in life.

Our findings are consistent with immunological alterations reported in adult MS patients. An increased $CD4^+/CD8^+$ ratio has been described in MS, as well as in other autoimmune disease patients and their relatives [6,7,32]. We now found that genetic predisposition for MS alters the $CD4^+/CD8^+$ ratio in children from the general population. This is relevant because the $CD4^+/CD8^+$ ratio is a relatively stable immunological parameter from childhood to young adulthood [33]. Our reported observed negative correlations of the MS-PRS with $CD8^+$ and not $CD4^+$ T cells are consistent with previously described reductions of $CD8^+$ and not $CD4^+$ T cells in the peripheral blood of MS patients [6]. In late stage adult MS, there is a local enrichment of memory $CD8^+$ T cells in the central nervous system (CNS) [34,35]. Accordingly, in MS-discordant monozygotic twins, a clonally expanded memory $CD8^+$ T cell population was found in the cerebrospinal fluid of twins asymptomatic for MS [36]. In children, however, it is not likely that genetically determined reductions in $CD8^+$ T cells are the indirect result of an increased influx of memory cells into the CNS, because both central and effector memory T cell populations did not correlate with the MS-PRS. Instead, genetic variation may impair thymic output of naive $CD8^+$ T cells, a phenomenon described in MS patients at an early age possibly before disease onset, which is supported by our observed negative association between the MS-PRS and naive $CD8^+$ T cells [6,37]. That the positive correlation of the MS-PRS with the $CD4^+/CD8^+$ ratio is mainly driven by major MHC risk alleles could imply that in children, the development of CD4 and CD8 double- into single-positive thymocytes is influenced by genetically altered HLA Class I- and II-expressing thymic epithelial cells [38,39].

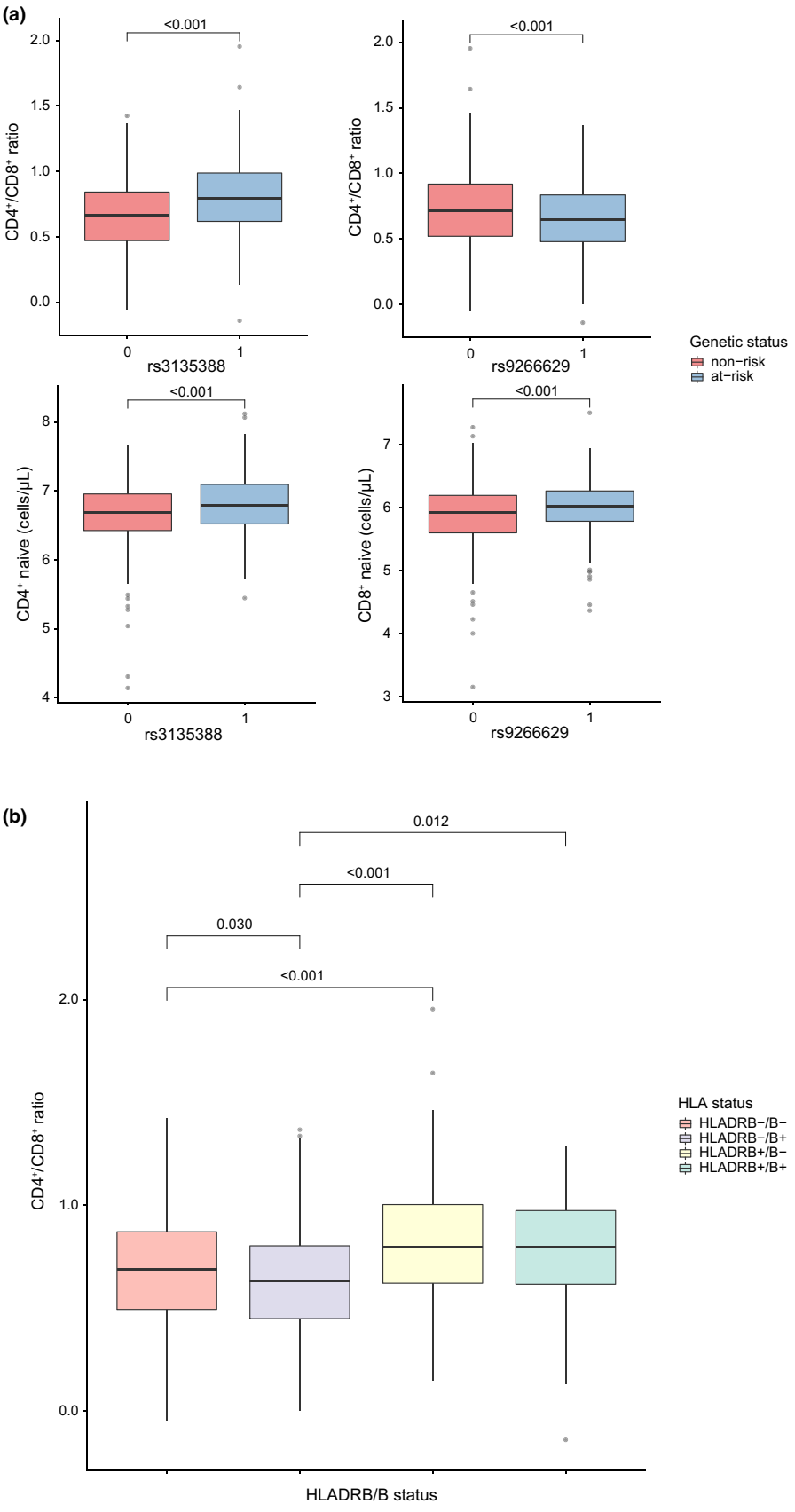

FIGURE 3 Associations of rs3135388 and rs9266629 with T cell subsets and the $CD4^+/CD8^+$ ratio. Carriership of rs3135388 and rs9266629 was dichotomized into two groups, homozygous non-risk and at-risk. At-risk participants consisted of heterozygous and homozygous carriers of the effect allele. (a) Left: Associations of rs3135388 with the $CD4^+/CD8^+$ ratio and $CD4^+$ naive cells. Right: Associations of rs9266629 with the $CD4^+/CD8^+$ ratio and $CD8^+$ naive cells. (b) Combined associations of rs3135388 and rs9266629 with the $CD4^+/CD8^+$ ratio [Colour figure can be viewed at wileyonlinelibrary.com]

The major risk variant for MS, *HLA-DRB1*15:01*, was found to be associated with higher $CD4^+$ T cell numbers, whereas the rs9266629 variant that is protective for MS and a tag-variant for *HLA-B* SNP rs3819284 associated positively with $CD8^+$ T cell numbers. The same associations were observed when incorporating these two variants in our MS-associated T cell PRSs. For $CD4^+$ T cells, it has

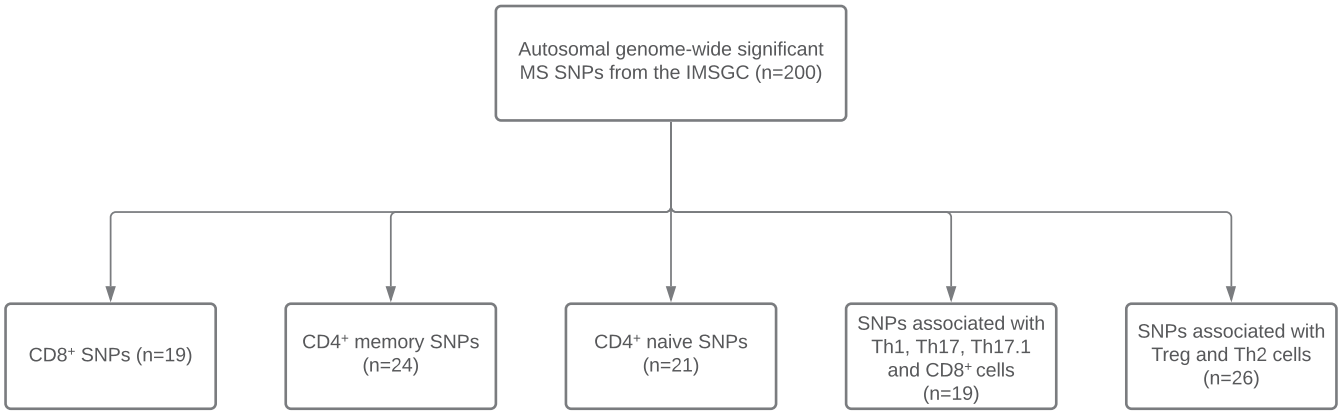

FIGURE 4 Flowchart describing the selection process of specific T cell polygenic risk scores using cis expression quantitative trait loci data. Risk variants were grouped using a combination of mRNA expression data, expression quantitative trait loci, and gene ontology analyses as described in the Materials and Methods section. IMSGC, International Multiple Sclerosis Genetics Consortium; MS, multiple sclerosis; SNP, single nucleotide polymorphism

been reported that naive populations are genetically affected and more activated and thus prone to differentiate into effector subsets during early MS [40]. Additionally, it was found that *HLA-DRB1*15:01* status increases autoproliferation and brain-homing of the $CD4^+$ T cells [41]. Disturbed T cell homeostasis in favor of $CD4^+$ T cells, influenced by rs3135388, the absence of rs9266629, and overall genetic risk for MS, may consequently be one of the first steps in the immunopathogenesis before possible onset of symptoms in MS. An increased proportion of these cells in peripheral blood could lay the foundation for the activation of autoreactive $CD4^+$ T cells later in life under the influence of environmental risk factors and *HLA-DRB1*15:01*, as suggested in the immunopathogenesis of MS [42]. Subsequent migration across the blood–brain barrier could initiate the MS disease process, where autoreactive T cells cause inflammatory demyelination of the CNS [42].

The entire MHC region contributed substantially to our results. Upon investigation of an MS-PRS with only MHC risk variants, our results were comparable to our PRS with all variants. Excluding the MHC region from the MS-PRS attenuated our results; however, they remained significant. This finding is in line with the widely known substantial role of the MHC region in MS pathogenesis. Due to this strong association of MHC alleles with MS (especially *HLA-DRB1*15:01*), we kept these variants in our PRSs, as this is the closest reflection of overall genetic MS risk; however, PRSs are not able to fully capture the effect of the MHC due to the epistatic interactions within this region of the genome.

We did not find an association between MS-PRSs and known environmental risk factors for MS (i.e., EBV and CMV serology and serum 25[OH]D). Low levels of serum 25(OH)D have been associated with increased MS risk in adults and children [43,44]. Our results suggest that serum 25(OH)D is not correlated with polygenic risk scores for MS at this age. This environmental risk factor could exert its effect on MS pathophysiology independently, a hypothesis also demonstrated by Mendelian randomization studies [45].

Persistent EBV and CMV seropositivity affects MS susceptibility in adults, but also in children [46,47]. The observed nonsignificant association in our study between polygenic risk for MS and EBV status could be explained by the EBV status assessment used in our study. EBV seropositivity in the Generation R study was tested by using EBV-VCA antibodies, which are not associated with MS in adults [48]. However, in pediatric MS, EBV-VCA positivity is significantly more present compared to controls, pleading for the validity of this marker in children [49]. Another possible explanation is that genetic MS risk does not associate with seropositivity for these viruses at this young age, and that EBV and CMV exert their effect at a later age in MS pathophysiology, as supported by earlier work [50].

Our study has several strengths. First, we were able to include a large number of young children, allowing us to detect robust effects of genetic MS risk even before possible disease onset. Furthermore, we incorporated several suggestive genetic risk variants for MS in our PRSs to reflect the polygenic architecture of MS as much as possible. Additionally, we tested for CMV positivity and serum 25(OH)D levels using accurate methods. Lastly, we performed additional analyses incorporating specific T cell PRSs constructed with functional pathway data.

Limitations are also present. Due to the cross-sectional design of our study, we are unable to investigate the temporal dynamics between genetic risk for MS and the T cell compartment of the immune system. Second, it was not possible to investigate the functional properties of T cells in this study, such as their proinflammatory and transmigration capacity. Future studies are needed to validate our results, test the reported relationships longitudinally, and incorporate environmental risk factors for MS to address their effect on T cells and to better understand the pathophysiology of MS. Longitudinal well-powered population-based study designs could potentially translate the altered $CD4^+/CD8^+$ ratio to a clinically usable biomarker for MS.

In conclusion, we report an association between genetic risk scores for MS and alterations in the T cell lineages in school-age children from the general population. This shows that MS genetics change the composition of the adaptive immune system during

childhood, which possibly contributes to overall risk of developing MS later in life.

ACKNOWLEDGMENTS

We are very thankful for the data provided by the International Multiple Sclerosis Genetics Consortium, which were used in constructing the polygenic risk scores for this study. We thank the children and parents, general practitioners, hospitals, midwives, and pharmacies in Rotterdam for their cooperation in the Generation R study.

CONFLICT OF INTEREST

R.F.N. participates in trials with Sanofi Genzyme and Novartis. None of the other authors has any conflict of interest to disclose.

AUTHOR CONTRIBUTIONS

Casper L. de Mol: Conceptualization (equal), formal analysis (equal), investigation (equal), methodology (equal), project administration (equal), visualization (equal), writing–original draft (equal), writing–review & editing (equal). **Kirsten I. M. Looman:** Conceptualization (equal), data curation (equal), formal analysis (equal), methodology (equal), writing–original draft (equal), writing–review & editing (equal). **Marvin M. van Luijn:** Conceptualization (equal), investigation (equal), methodology (equal), supervision (equal), writing–original draft (equal), writing–review & editing (equal). **Karim L. Kreft:** Conceptualization (equal), formal analysis (equal), investigation (equal), methodology (equal), writing–original draft (equal), writing–review & editing (equal). **Philip R. Jansen:** Conceptualization (equal), formal analysis (equal), investigation (equal), methodology (equal), writing–review & editing (equal). **Menno C. van Zelm:** Conceptualization (equal), data curation (equal), investigation (equal), methodology (equal), writing–review & editing (equal). **Joost Smolders:** Conceptualization (equal), methodology (equal), writing–original draft (equal), writing–review & editing (equal). **Tonya White:** Conceptualization (equal), methodology (equal), supervision (equal), writing–review & editing (equal). **Henriette A. Moll:** Conceptualization (equal), data curation (equal), formal analysis (equal), investigation (equal), methodology (equal), supervision (equal), writing–review & editing (equal). **Rinze F. Neuteboom:** Conceptualization (equal), formal analysis (equal), funding acquisition (equal), investigation (equal), methodology (equal), resources (equal), supervision (equal), writing–original draft (equal), writing–review & editing (equal).

DATA AVAILABILITY STATEMENT

The data used in this study are not publicly available due to legal and informed consent restrictions. However, qualified researchers can request to access the data by contacting the Generation R study (datamanagementgenr@erasmusmc.nl).

ORCID

Casper L. de Mol https://orcid.org/0000-0002-3733-1706
Joost J. F. M. Smolders https://orcid.org/0000-0001-9766-8661

REFERENCES

1. McLeod JG, Hammond SR, Kurtzke JF. Migration and multiple sclerosis in immigrants to Australia from United Kingdom and Ireland: a reassessment. I. Risk of MS by age at immigration. *J Neurol*. 2011;258(6):1140-1149. https://doi.org/10.1007/s00415-010-5898-4
2. Olsson T, Barcellos LF, Alfredsson L. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nat Rev Neurol*. 2016;13(1):26-36. https://doi.org/10.1038/nrneurol.2016.187
3. Consortium IMSG. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science*. 2019;365(6460):eaav7188. https://doi.org/10.1126/science.aav7188
4. O'Gorman C, Lin R, Stankovich J, Broadley SA. Modelling genetic susceptibility to multiple sclerosis with family data. *Neuroepidemiology*. 2012;40(1):1-12. https://doi.org/10.1159/000341902
5. Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet*. 2013;9(3):e1003348. https://doi.org/10.1371/journal.pgen.1003348
6. Pender MP, Csurhes PA, Pfluger CMM, Burrows SR. Deficiency of CD8+ effector memory T cells is an early and persistent feature of multiple sclerosis. *Mult Scler J*. 2014;20(14):1825-1832. https://doi.org/10.1177/1352458514536252
7. Stüve O, Marra CM, Bar-Or A, et al. Altered CD4+/CD8+ T-cell ratios in cerebrospinal fluid of natalizumab-treated patients with multiple sclerosis. *Arch Neurol*. 2006;63(10):1383-1387. https://doi.org/10.1001/archneur.63.10.1383
8. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. *Eur J Epidemiol*. 2016;31(12):1243-1264. https://doi.org/10.1007/s10654-016-0224-9
9. Van Den Heuvel D, Jansen MAE, Dik WA, et al. Cytomegalovirus- and epstein-barr virus-induced T-cell expansions in young children do not impair naive T-cell populations or vaccination responses: the Generation R study. *J Infect Dis*. 2016;213(2):233-242. https://doi.org/10.1093/infdis/jiv369
10. Medina-Gomez C, Felix JF, Estrada K, et al. Challenges in conducting genome-wide association studies in highly admixed multi-ethnic populations: the Generation R Study. *Eur J Epidemiol*. 2015;30(4):317-330. https://doi.org/10.1007/s10654-015-9998-4
11. Jansen PR, Muetzel RL, Polderman TJC, et al. Polygenic scores for neuropsychiatric traits and white matter microstructure in the pediatric population. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2019;4(3):243-250. https://doi.org/10.1016/j.bpsc.2018.07.010
12. International HapMap Consortium. The International HapMap project. *Nature*. 2003;426(6968):789-796. https://doi.org/10.1038/nature02168
13. University of Michigan 1000G Phase I Integrated Release Version 3 Haplotypes. Published 2012. http://csg.sph.umich.edu/abecasis/MACH/download/1000G.2012-03-14.html
14. Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. *Bioinformatics*. 2015;31(9):1466-1468. https://doi.org/10.1093/bioinformatics/btu848
15. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575. https://doi.org/10.1086/519795
16. Živković M, Stanković A, Dinčić E, et al. The tag SNP for HLA-DRB1*1501, rs3135388, is significantly associated with multiple sclerosis susceptibility: cost-effective high-throughput detection by real-time PCR. *Clin Chim Acta*. 2009;406(1–2):27-30. https://doi.org/10.1016/j.cca.2009.05.004
17. Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*. 2014;506(7488):376-381. https://doi.org/10.1038/nature12873.

18. De Bakker PIW, McVean G, Sabeti PC, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet*. 2006;38(10):1166-1172. https://doi.org/10.1038/ng1885
19. Monaco G, Lee B, Xu W, et al. RNA-Seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Rep*. 2019;26(6):1627-1640.e7. https://doi.org/10.1016/j.celrep.2019.01.041
20. Schmiedel BJ, Singh D, Madrigal A, et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell*. 2018;175(6):1701-1715.e16. https://doi.org/10.1016/j.cell.2018.10.022
21. Chandra V, Bhattacharyya S, Schmiedel BJ, et al. Promoter-interacting expression quantitative trait loci are enriched for functional genetic variants. *Nat Genet*. 2021;53(1):110-119. https://doi.org/10.1038/s41588-020-00745-3
22. Wu C, Jin X, Tsueng G, Afrasiabi C, Su AI. BioGPS: building your own mash-up of gene annotations and expression profiles. *Nucleic Acids Res*. 2016;44(D1):D313-D316. https://doi.org/10.1093/nar/gkv1104
23. Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res*. 2012;41(D1):D377-D386. https://doi.org/10.1093/nar/gks1118
24. Healy J, Thomas EE, Schwartz JT, Wigler M. Annotating large genomes with exact word matches. *Genome Res*. 2003;13(10):2306-2315. https://doi.org/10.1101/gr.1350803
25. Huerta-Cepas J, Szklarczyk D, Heller D, et al. EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*. 2019;47(D1):D309-D314. https://doi.org/10.1093/nar/gky1085
26. van der Tas JT, Elfrink MEC, Heijboer AC, et al. Foetal, neonatal and child vitamin D status and enamel hypomineralization. *Community Dent Oral Epidemiol*. 2018;46(4):343-351. https://doi.org/10.1111/cdoe.12372
27. Kalina T, Flores-Montero J, Van Der Velden VHJ, et al. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. *Leukemia*. 2012;26(9):1986-2010. https://doi.org/10.1038/leu.2012.122
28. Pedreira CE, Costa ES, Lecrevisse Q, van Dongen JJM, Orfao A. Overview of clinical flow cytometry data analysis: recent advances and future challenges. *Trends Biotechnol*. 2013;31(7):415-425. https://doi.org/10.1016/j.tibtech.2013.04.008
29. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/. Accessed April 7, 2021.
30. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 1995;57(1):289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
31. Sinotte M, Diorio C, Bérubé S, Pollak M, Brisson J. Genetic polymorphisms of the vitamin D binding protein and plasma concentrations of 25-hydroxyvitamin D in premenopausal women. *Am J Clin Nutr*. 2009;89(2):634-640. https://doi.org/10.3945/ajcn.2008.26445
32. Pender MP. CD8+ t-cell deficiency, epstein-barr virus infection, vitamin d deficiency, and steps to autoimmunity: a unifying hypothesis. *Autoimmune Dis*. 2012;2012:1-16. https://doi.org/10.1155/2012/189096
33. Comans-Bitter WM, De Groot R, Van den Beemd R, et al. Immunophenotyping of blood lymphocytes in childhood: reference values for lymphocyte subpopulations. *J Pediatr*. 1997;130(3):388-393. https://doi.org/10.1016/S0022-3476(97)70200-2
34. Fransen NL, Hsiao CC, Van Der Poel M, et al. Tissue-resident memory T cells invade the brain parenchyma in multiple sclerosis white matter lesions. *Brain*. 2020;143(6):1714-1730. https://doi.org/10.1093/brain/awaa117
35. Machado-Santos J, Saji E, Tröscher AR, et al. The compartmentalized inflammatory response in the multiple sclerosis brain is composed of tissue-resident CD8+ T lymphocytes and B cells. *Brain*. 2018;141(7):2066-2082. https://doi.org/10.1093/brain/awy151
36. Beltrán E, Gerdes LA, Hansen J, et al. Early adaptive immune activation detected in monozygotic twins with prodromal multiple sclerosis. *J Clin Invest*. 2019;129(11):4758-4768. https://doi.org/10.1172/JCI128475
37. Duszczyszyn DA, Beck JD, Antel J, et al. Altered naive CD4 and CD8 T cell homeostasis in patients with relapsing-remitting multiple sclerosis: thymic versus peripheral (non-thymic) mechanisms. *Clin Exp Immunol*. 2006;143(2):305-313. https://doi.org/10.1111/j.1365-2249.2005.02990.x
38. Tsai S, Santamaria P. MHC class II polymorphisms, autoreactive T-cells, and autoimmunity. *Front Immunol*. 2013;4:321. https://doi.org/10.3389/fimmu.2013.00321
39. Kisielow P, Teh HS, Blüthmann H, Von Boehmer H. Positive selection of antigen-specific T cells in thymus by restricting MHC molecules. *Nature*. 1988;335(6192):730-733. https://doi.org/10.1038/335730a0
40. Corvol JC, Pelletier D, Henry RG, et al. Abrogation of T cell quiescence characterizes patients at high risk for multiple sclerosis after the initial neurological event. *Proc Natl Acad Sci USA*. 2008;105(33):11839-11844. https://doi.org/10.1073/pnas.0805065105
41. Jelcic I, Al Nimer F, Wang J, et al. Memory B cells activate brain-homing, autoreactive CD4+ T cells in multiple sclerosis. *Cell*. 2018;175(1):85-100.e23. https://doi.org/10.1016/j.cell.2018.08.011
42. Baecher-Allan C, Kaskow BJ, Weiner HL. Multiple sclerosis: mechanisms and immunotherapy. *Neuron*. 2018;97(4):742-768. https://doi.org/10.1016/j.neuron.2018.01.021
43. Munger KL, Levin LI, Hollis BW, Howard NS, Ascherio A. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. *J Am Med Assoc*. 2006;296(23):2832-2838. https://doi.org/10.1001/jama.296.23.2832
44. Banwell B, Bar-Or A, Arnold DL, et al. Clinical, environmental, and genetic determinants of multiple sclerosis in children with acute demyelination: a prospective national cohort study. *Lancet Neurol*. 2011;10(5):436-445. https://doi.org/10.1016/S1474-4422(11)70045-X
45. Belbasis L, Bellou V, Evangelou E, Tzoulaki I. Environmental factors and risk of multiple sclerosis: findings from meta-analyses and Mendelian randomization studies. *Mult Scler J*. 2020;26(4):397-404. https://doi.org/10.1177/1352458519872664
46. Maple PAC, Tanasescu R, Gran B, Constantinescu CS. A different response to cytomegalovirus (CMV) and Epstein-Barr virus (EBV) infection in UK people with multiple sclerosis (PwMS) compared to controls. *J Infect*. 2020;80(3):320-325. https://doi.org/10.1016/j.jinf.2019.10.017
47. Waubant E, Mowry EM, Krupp L, et al. Common viruses associated with lower pediatric multiple sclerosis risk. *Neurology*. 2011;76(23):1989-1995. https://doi.org/10.1212/WNL.0b013e31821e552a
48. Lünemann JD, Tintoré M, Messmer B, et al. Elevated Epstein-Barr virus-encoded nuclear antigen-1 immune responses predict conversion to multiple sclerosis. *Ann Neurol*. 2010;67(2):159-169. https://doi.org/10.1002/ana.21886
49. Waubant E, Mowry EM, Krupp L, et al. Common viruses associated with lower pediatric multiple sclerosis risk. *Neurology*. 2011;76(23):1989-1995. https://doi.org/10.1212/WNL.0b013e31821e552a
50. Hedström AK, Huang J, Michel A, et al. High levels of Epstein-Barr virus nuclear antigen-1-specific antibodies and infectious mononucleosis act both independently and synergistically to increase

multiple sclerosis risk. *Front Neurol*. 2020;10:1368. https://doi.org/10.3389/fneur.2019.01368

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: de Mol CL, Looman KIM, van Luijn MM, et al. T cell composition and polygenic multiple sclerosis risk: A population-based study in children. *Eur J Neurol*. 2021;28:3731–3741. https://doi.org/10.1111/ene.15019

REGISTER NOW

Rare Disease Connect in Neurology

10–12 November 2021

Three half-days of virtual collaboration

04:00–08:00 PST (Los Angeles) I 07:00–11:00 EST (New York)
12:00–16:00 GMT (London) I 13:00–17:00 CET (Paris)
19:00–23:00 JST (Tokyo)

Join us for **Rare Disease Connect in Neurology 2021**, the inaugural meeting of a long-term learning journey, connecting neuromuscular experts committed to delivering excellence for patients living with myasthenia gravis (MG).

Imagine what we could do for patients, together.

Why take part in Rare Disease Connect in Neurology 2021?

With a world-class programme spanning three half-days, developed in collaboration with a steering committee of renowned MG experts, **Rare Disease Connect in Neurology 2021** will feature:

- Live and on-demand content
- Group discussions
- Networking opportunities with your global peers

To join us virtually on 10–12 November 2021

Register your interest now

Inspired by **patients.**
Driven by **science.**