Vitiligo age-of-onset and PRS Association

Genevieve HL Roberts

2024-4-28

Introduction

In this study, we aim to investigate the association between a vitiligo Polygenic Risk Score (PRS) excluding MHC Class II region SNPs and the age-of-onset. Our hypothesis posits a correlation wherein a higher PRS (excluding MHC Class II) is linked to an earlier age-of-onset, indicating a potential association with elevated genetic risk.

Moreover, we extend our hypothesis to carriers of the extreme risk/early age of onset MHC Class II haplotype. In this subgroup, we anticipate that the impact of a high PRS without MHC Class II may be diminished due to the substantial genetic risk conferred by the haplotype (haplotype OR=8.1). We believe there may be an interaction effect, suggesting that the slope of the regression line for age-of-onset ~ PRS is smaller in individuals carrying the high-risk haplotype compared to non-carriers.

Results

Categorizing age-of-onset into Early- and Late- Onset Groups

I no longer have the AOO classifications that we created previously, so I have re-fit the finite mixture model to categorize GWAS123 cases as either early onset or late onset. I have used the same methods here.

In the eAOO paper it says, "After subgroup assignment and control matching, the combined early-onset subgroup contained 704 cases and 9,031 controls and the combined late-onset subgroup contained 1,467 cases and 19,156 controls."

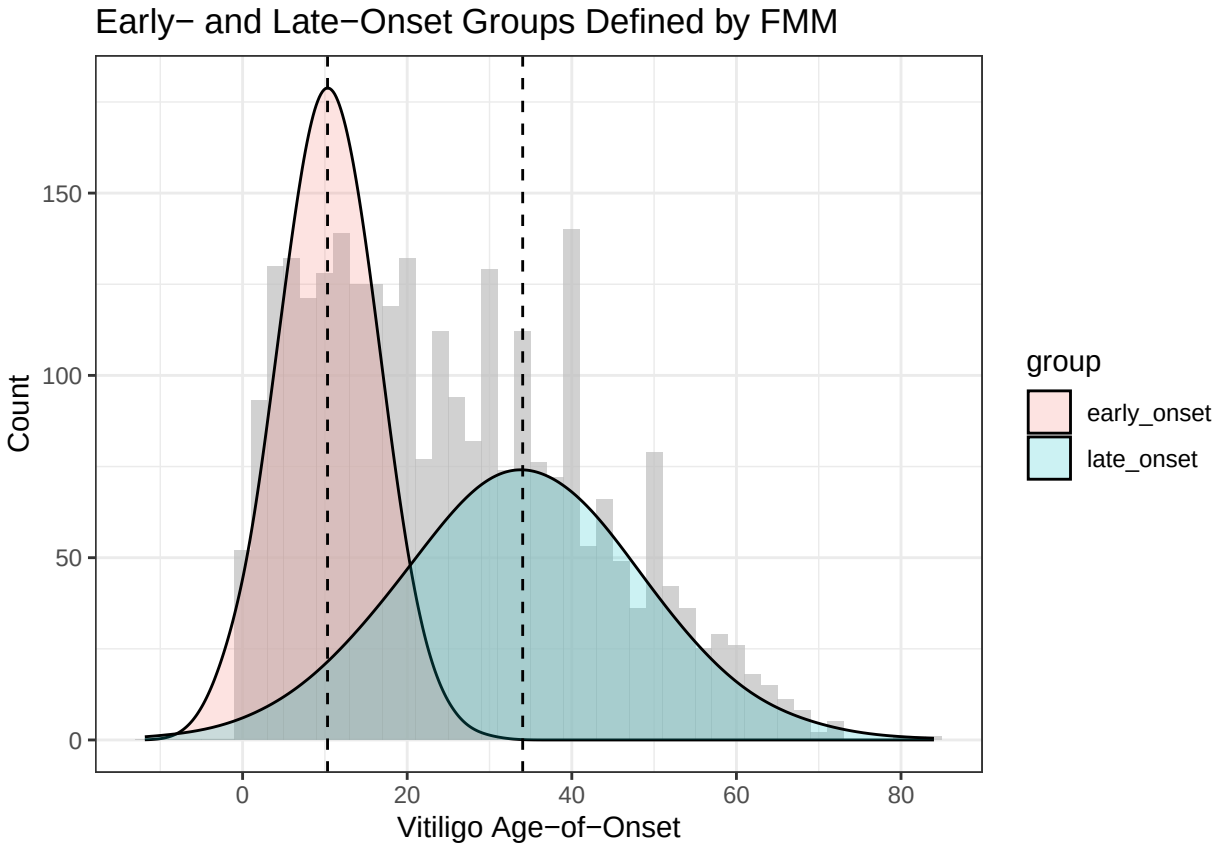
number of iterations= 100

Table 1: Early and Late Onset Subgroup Descriptive Statistics (continued below)

Age of Onset Category	n	FMM Mean	Actual Mean	FMM SD	Actual SD
early_onset	719	10.32	6.68	5.86	3.39
late_onset	1431	34.03	39.08	14.4	11.4
NA	617	NA	17.19	NA	2.82

Actual Minimum	Actual Maximum
0	12
23	84
12.5	22

As shown in the table above, you can see that we achieve similar sample sizes for early and late onset (719 early here vs. 704 in our earlier analysis and 1,431 late here vs. 1,467 in our earlier analysis).The FMM-estimated distribution means are also identical, at 10.3 here versus 10.3 for GWAS123 + Rep in our earlier analysis and 34.0 here versus 34.0 for GWAS123 + Rep in our earlier analysis.


Early− and Late−Onset Groups Defined by FMM
Count
150
100
50
0
0
20
40
60
80
Vitiligo Age−of−Onset
group
early_onset
late_onset

PRS distributions

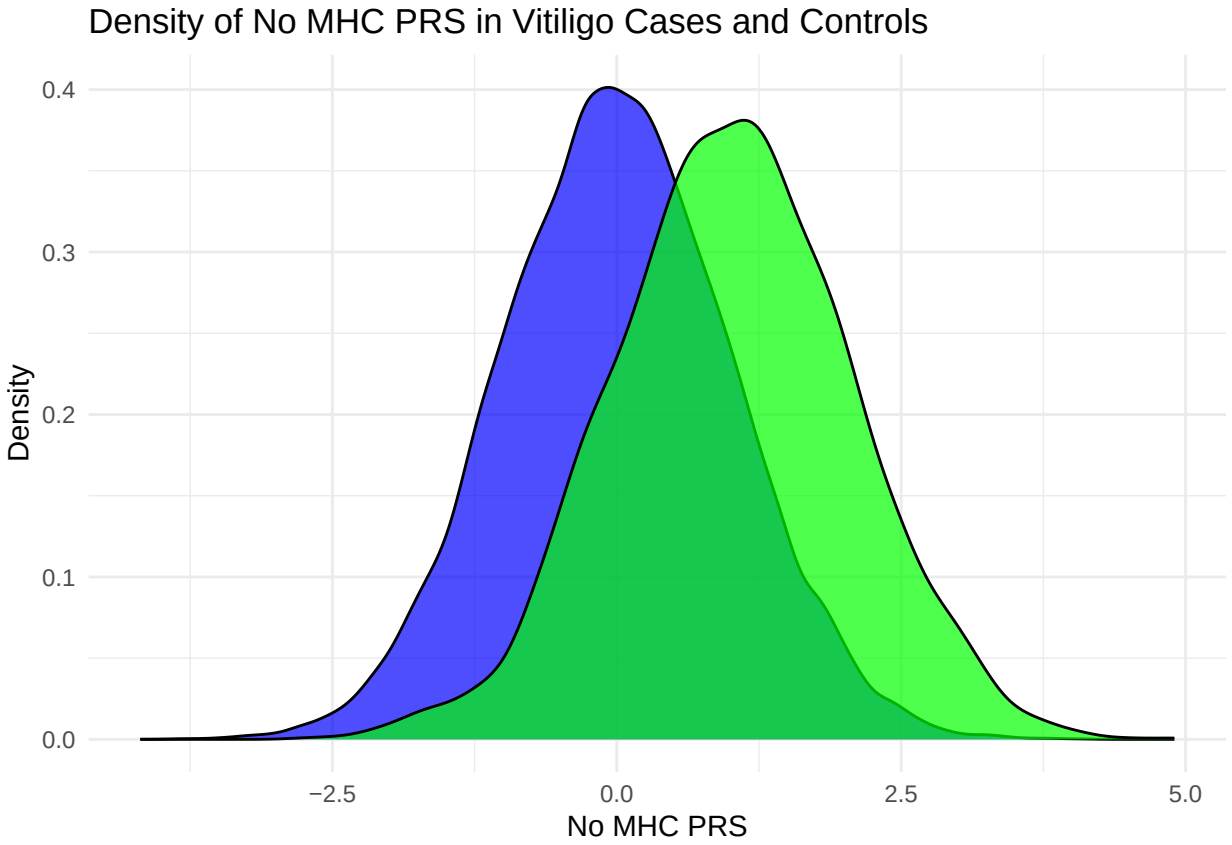


Table 3: Table continues below

age_onset_category	no_mhc_classII_prs_category	n	mean_no_mhc_prs
Early Onset	High PRS	116	2.573
Early Onset	Normal PRS	603	0.6777
Late Onset	High PRS	234	2.533
Late Onset	Normal PRS	1197	0.6714

mean_no_mhc_percentile	sum	percent
0.9911	719	0.1613
0.7076	719	0.8387
0.9907	1431	0.1635
0.7045	1431	0.8365

Associations with Vitiligo Risk (Case-Control)

Next, I wanted to verify that the associations looked as I would expect for each of the PRSs. So, I checked the association with overall vitiligo risk (i.e. in all cases and controls in GWAS123):

Phenotype	PRS or SNP	pval	OR	lower_ci	upper_ci
vitiligo	CONFIRMED	0.00e+00	2.71	2.6	2.82
vitiligo	eA00_rs145954018	1.74e-49	2.35	2.1	2.64
vitiligo	eAOO_hap_carrier	2.33e-49	2.43	2.16	2.74
vitiligo	generic_rs9271597	1.90e-97	1.84	1.74	1.95
vitiligo	mhc_class2_only	6.12e-110	1.46	1.42	1.51
vitiligo	no_mhc_classII	0.00e+00	2.59	2.48	2.7
late_onset_vitiligo	CONFIRMED	3.61e-230	2.46	2.33	2.6
late_onset_vitiligo	eA00_rs145954018	9.05e-08	1.61	1.35	1.92
late_onset_vitiligo	eAOO_hap_carrier	2.22e-07	1.62	1.35	1.94
late_onset_vitiligo	generic_rs9271597	1.26e-31	1.59	1.47	1.71
late_onset_vitiligo	mhc_class2_only	1.08e-27	1.3	1.24	1.36
late_onset_vitiligo	no_mhc_classII	2.25e-222	2.51	2.37	2.66
early_onset_vitiligo	CONFIRMED	4.93e-192	3.06	2.84	3.3
early_onset_vitiligo	eA00_rs145954018	2.38e-75	4.7	3.99	5.55
early_onset_vitiligo	eAOO_hap_carrier	4.30e-75	5.14	4.32	6.13
early_onset_vitiligo	generic_rs9271597	4.00e-51	2.31	2.08	2.58
early_onset_vitiligo	mhc_class2_only	4.99e-92	1.82	1.72	1.93
early_onset_vitiligo	no_mhc_classII	1.61e-127	2.55	2.36	2.75

These results look generally as I would expect, with the CONFIRMED risk score having the highest performance with respect to OR per standard deviation of PRS. The P-values = 0 mean that the P-value is lower than the numerical precision in R, which I believe is something like P < 1e-300.

Note that the scaling is different for SNPs and for PRS, so the estimates are not directly comparable (for PRS, the OR is per standard deviation of PRS, whereas for SNPs, it is per SNP risk allele).

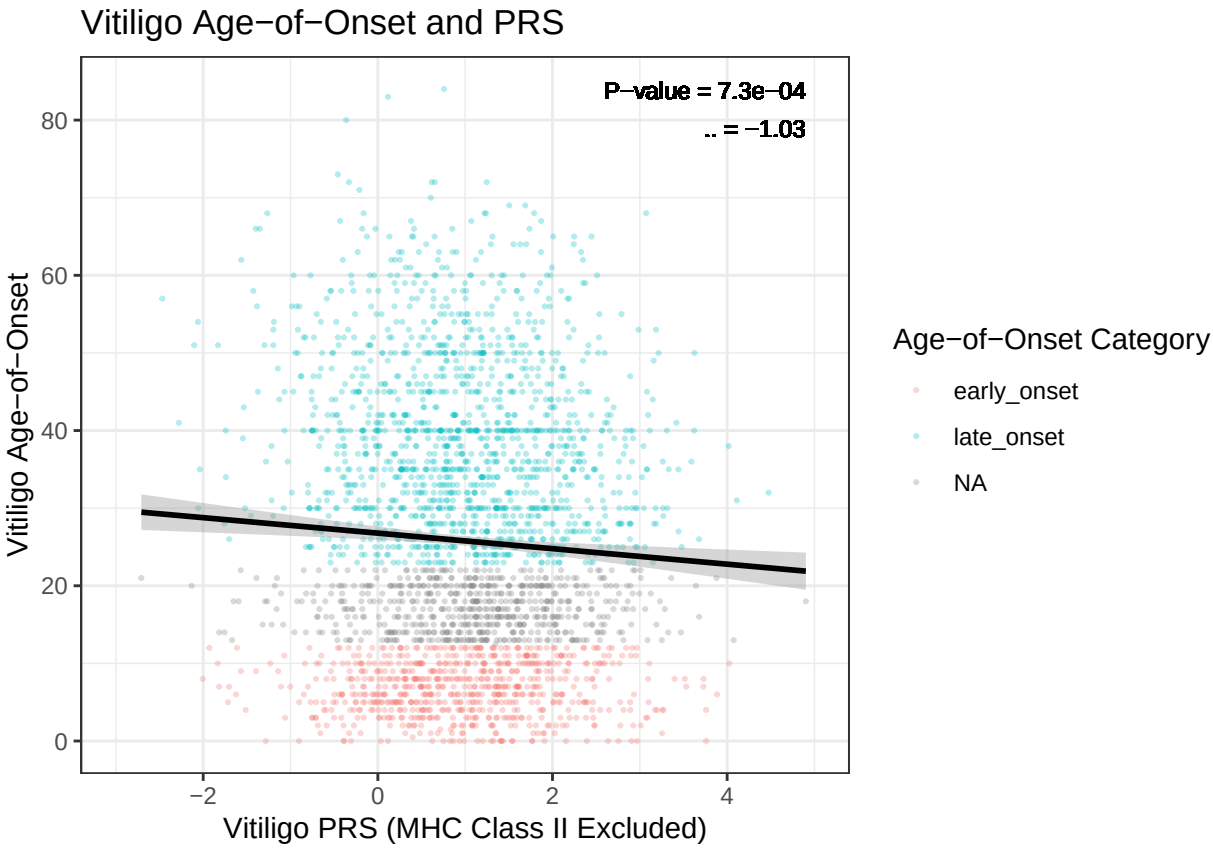
Associations with Vitiligo Age-of-Onset (Case Only)

I also checked the association with vitiligo age-of-onset risk (i.e. cases with non-missing age-of-onset in GWAS123):

PRS or SNP	pval	estimate	lower_ci	upper_ci	n_obs	adj_r_sq
no_mhc_classII_prs	7.30e-04	-1.03	-1.63	-0.43	2766	0.00701
mhc_class2_only_prs	2.94e-22	-2.59	-3.11	-2.07	2767	0.0364
CONFIRMED_prs	1.59e-15	-2.37	-2.96	-1.79	2766	0.0257
eA00_rs145954018	9.74e-18	-7.29	-8.94	-5.64	2767	0.0292
generic_rs9271597	6.32e-11	-2.91	-3.78	-2.04	2767	0.0182
eAOO_hap_carrier	3.12e-18	-7.71	-9.44	-5.99	2766	0.03

Again, these results look about as I would expect: the MHC Class II only PRS is more strongly associated than the confirmed PRS. Nevertheless, there is modest association between the non-MHC Class II PRS and lower vitiligo age-of-onset, as we hypothesized (P=7.30e-04).

Note that the scaling is different for SNPs and for PRS, so the estimates are not directly comparable (for PRS, the estimate is interpreted as years per standard deviation of PRS, whereas for SNPs, it is interpreted as years per copy of SNP risk allele).

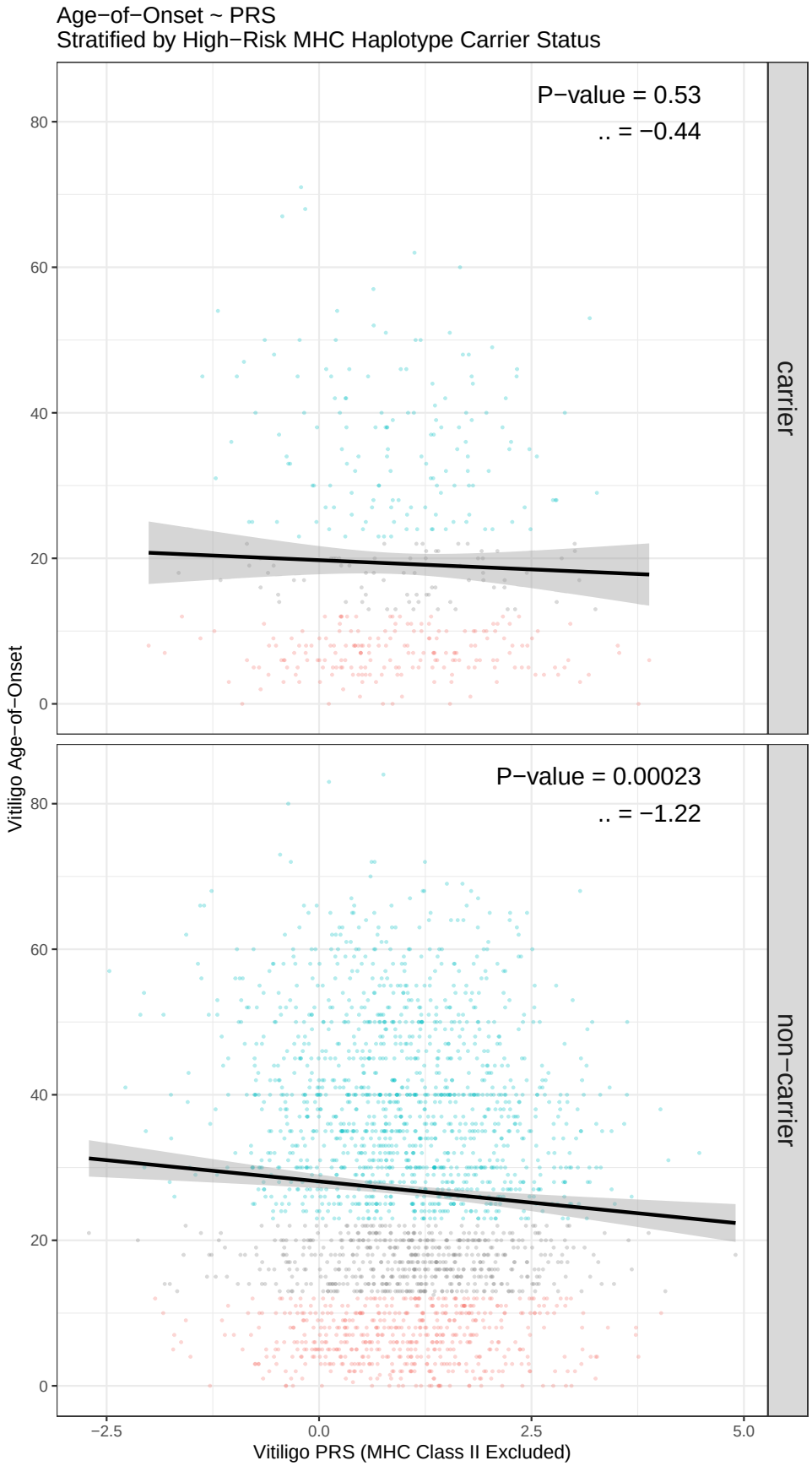


Stratified Association by High-Risk MHC Haplotype Carrier Status

Table 7: Stratified Association between PRS and Age-of-Onset in High Risk Haplotype Carrier Groups (continued below)

high_risk_MHC_haplotype_carrier	Phenotype	PRS	pval
carrier	VITageonset	no_mhc_classII	0.53
non-carrier	VITageonset	no_mhc_classII	0.00023

estimate	lower_ci	upper_ci	n_obs	adj_r_sq
-0.44	-1.82	0.94	408	0.0243
-1.22	-1.87	-0.57	2357	0.00827


Age−of−Onset ~ PRS
Stratified by High−Risk MHC Haplotype Carrier Status
P−value = 0.53
.. = −0.44
carrier
P−value = 0.00023
.. = −1.22
non−carrier
Vitiligo Age−of−Onset
Vitiligo PRS (MHC Class II Excluded)
0
20
40
60
80
−2.5
0.0
2.5
5.0

AOO_category	high_risk_MHC_haplotype_carrier	n	sum	percent
early_onset	carrier	190	718	26.46
early_onset	non-carrier	528	718	73.54
late_onset	carrier	145	1431	10.13
late_onset	non-carrier	1286	1431	89.87

Stratified Association in Early and Late Onset Groups

Table 10: Stratified Association between PRS and Age-of-Onset in Early- and Late-Onset Groups (continued below)

AOO_category	Phenotype	PRS	pval	estimate	lower_ci
early_onset	VITageonset	no_mhc_classII	0.61	0.08	-0.22
late_onset	VITageonset	no_mhc_classII	5.5e-06	-1.4	-2

upper_ci	n_obs	adj_r_sq
0.38	528	0.00867
-0.8	1285	0.0182

Look at assocaition of PRS in MHC High Risk Haplotype Carriers

Table 12: Table continues below

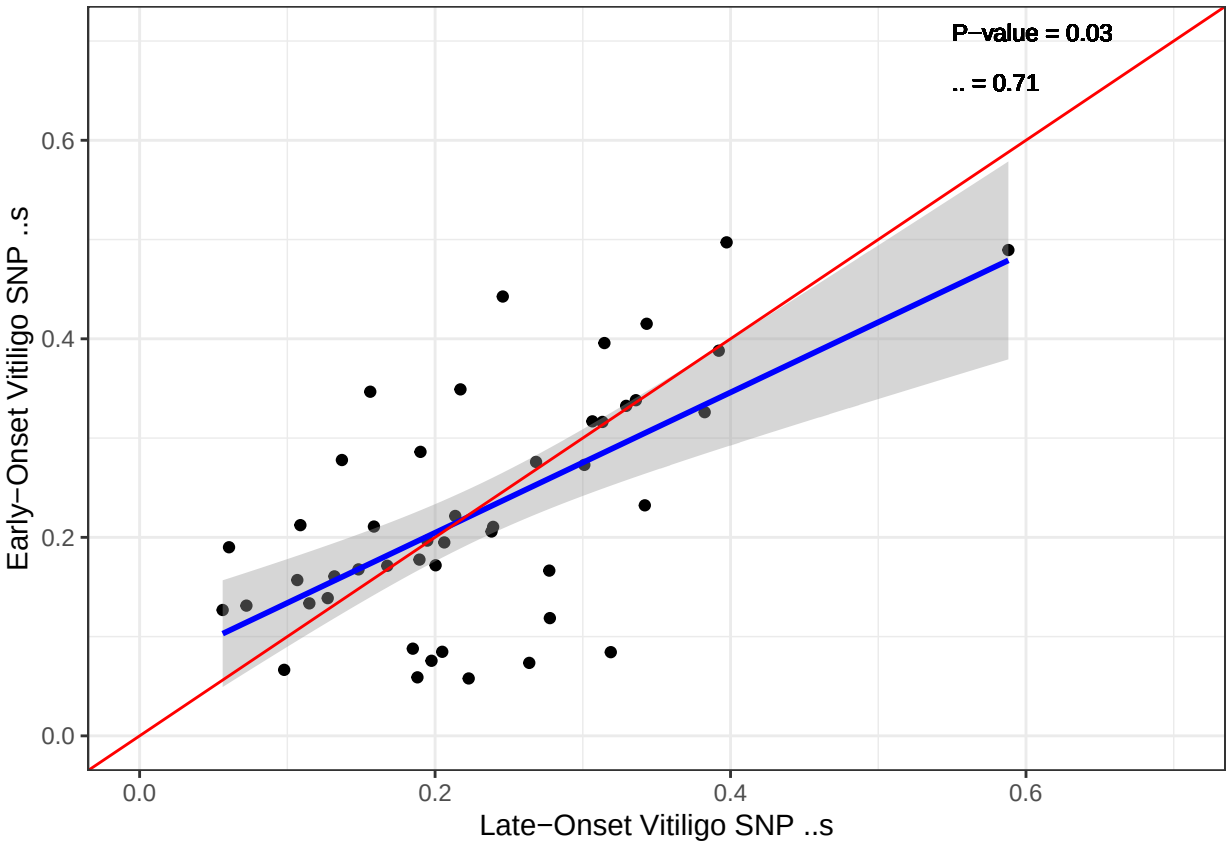
Phenotype	PRS	term	pval
vitiligo	no_mhc_classII_prs_scaled	no_mhc_classII_prs_scaled	2.2e-42

estimate	std_error	OR	lower_ci	upper_ci
0.84	0.06	2.32	2.06	2.62

Compare non-MHC SNP effect estimates in early- and late-onset

First, I have downsampled cases so that there are exactly 700 cases with early-onset and 700 cases with late-onset. Then, I computed the association between each of the non-MHC lead GWAS variants and disease risk.

Below, I have plotted the effect estimates in the early-onset and late-onset group together. If the other loci have generally equivalent effects in both groups, we expect that the points should generally fall around the y=x line (denoted in red). If, on the other hand, there is effect size dilution in the early-onset group, we expect that the slope will be <1.


P−value = 0.03
.. = 0.71
Early−Onset Vitiligo SNP ..s
Late−Onset Vitiligo SNP ..s
0.0
0.2
0.4
0.6

Year of Onset Analysis

For a subset of cases, I was able to find year-off-onset data.

Table 14: Table continues below

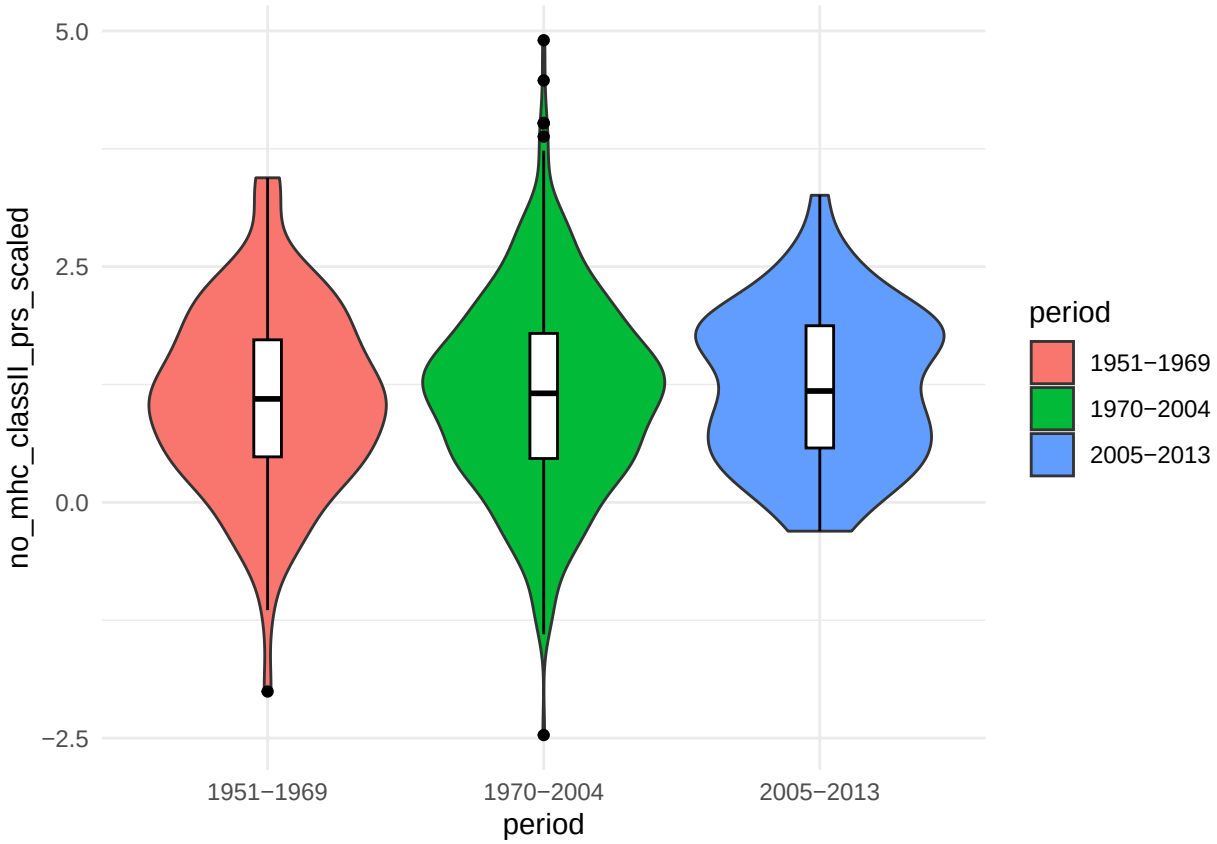
Period	n Total Vitiligo Cases	Average Prob. of Early Onset
1951-1969	108	0.656
1970-2004	537	0.3707
2005-2013	56	0.276

n MHC High Risk Haplotype Carriers	% MHC High Risk Haplotype Carriers	Mean no-MHC PRS
13	12.04	1.095
68	12.66	1.144
7	12.5	1.229

Association between eAOO SNP and year of onset

term	estimate	std.error	statistic	p.value
(Intercept)	1987	1.274	1559	0
eAOO_hap_carrier	-0.07936	1.629	-0.04872	0.9612
g1p1	79.65	81.77	0.9741	0.3304
g1p2	-26.77	92.32	-0.29	0.7719
g2p1	124.5	117.5	1.059	0.2899
g2p2	77.33	90.77	0.852	0.3945
g2p3	-166.8	222.5	-0.7497	0.4537
g2p4	72.89	117.9	0.6183	0.5366
g3p1	305.9	235.9	1.297	0.1952
g3p2	-313	175.7	-1.781	0.0754
g3p3	124.3	254.8	0.4878	0.6259
g3p4	154	202.4	0.7607	0.4471
g3p5	181.9	158.8	1.145	0.2524
g3p6	241.5	128.4	1.881	0.06036
g3p7	-17.01	158.9	-0.1071	0.9148
g3p8	157.6	138.8	1.136	0.2566
sex	-2.375	1.357	-1.751	0.08045

Association between PRS and year of onset period We noticed that in the table, the mean PRS appears as though it may be increasing with year of onset. We explicitly test that hypothesis here.



term	estimate	std.error	statistic	p.value
(Intercept)	1.115	0.1162	9.6	1.474e-20
period_ordinal	0.03142	0.07898	0.3978	0.6909
g1p1	2.426	5.695	0.426	0.6703
g1p2	-2.837	6.418	-0.442	0.6586
g2p1	-2.381	8.188	-0.2908	0.7713
g2p2	6.826	6.328	1.079	0.2811
g2p3	-25.24	15.5	-1.628	0.104
g2p4	-0.3657	8.211	-0.04454	0.9645
g3p1	17.23	16.45	1.048	0.295
g3p2	-30.15	12.24	-2.463	0.01404
g3p3	-6.093	17.62	-0.3458	0.7296
g3p4	12.15	14.05	0.8644	0.3877
g3p5	8.405	11.06	0.7601	0.4474
g3p6	3.014	8.958	0.3365	0.7366
g3p7	0.3431	10.99	0.0312	0.9751
g3p8	2.691	9.666	0.2784	0.7808
sex	-0.04222	0.09437	-0.4474	0.6547