

1 **Autoimmune diseases and risk of non-Hodgkin lymphoma: A Mendelian randomisation**
2 **study**

3 Xiaoting Shi, PhD Candidate^{1,4}; Joshua D. Wallach, Assistant Professor²; Xiaomei Ma,
4 Professor³; Tormod Rogne, Assistant Professor^{3,4}

5 ¹ Department of Environmental Health Sciences, Yale School of Public Health, New Haven,
6 Connecticut, USA

7 ² Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta,
8 Georgia, USA

9 ³ Department of Chronic Diseases Epidemiology, Yale School of Public Health, New Haven,
10 Connecticut, USA

11 ⁴ Yale Center for Perinatal, Pediatric, and Environmental Epidemiology, Yale School of Public
12 Health, New Haven, Connecticut, USA

13 **Corresponding author:**

14 Xiaoting Shi, MPhil

15 Department of Environmental Health Sciences, Yale School of Public Health, New Haven,
16 Connecticut, USA

17 xiaoting.shi@yale.edu

18 ORCID number: Xiaoting Shi <https://orcid.org/0000-0002-5256-0646>

19 Joshua Wallach <https://orcid.org/0000-0002-2816-6905>

20 Tormod Rogne: <https://orcid.org/0000-0002-9581-7384>

21 Twitter: @Xiaoting_yale, @JoshuaDWallach, @TormodRogne

22 Word count: 3,394

23 Figures/Tables: 3/2

24 References: 56

25 Appendix: 2

26

27

28 **ABSTRACT**

29 **Objective:** To examine whether genetically predicted susceptibility to ten autoimmune diseases
30 (Behçet's disease, coeliac disease, dermatitis herpetiformis, lupus, psoriasis, rheumatoid arthritis,
31 sarcoidosis, Sjögren's syndrome, systemic sclerosis, and type 1 diabetes) is associated with risk
32 of non-Hodgkin lymphoma (NHL).

33 **Design:** Two sample Mendelian randomization (MR) study.

34 **Setting:** Genome wide association studies (GWASs) of ten autoimmune diseases, NHL, and four
35 NHL subtypes (i.e., follicular lymphoma, mature T/natural killer-cell lymphomas, non-follicular
36 lymphoma, and other and unspecified types of NHL).

37 **Analysis** We used data from the largest publicly available GWASs of European ancestry for
38 each autoimmune disease, NHL, and NHL subtypes. For each autoimmune disease, we extracted
39 single nucleotide polymorphisms (SNPs) strongly associated ($P < 5 \times 10^{-8}$) with that disease and
40 that were independent of one another ($R^2 < 1 \times 10^{-3}$) as genetic instruments. SNPs within the
41 human leukocyte antigen region were not considered due to potential pleiotropy. Our primary
42 MR analysis was the inverse-variance weighted analysis. Additionally, we conducted MR-Egger,
43 weighted mode, and weighted median regression to address potential bias due to pleiotropy, and
44 robust adjusted profile scores to address weak instrument bias. We carried out sensitivity
45 analysis limited to the non-immune pathway for nominally significant findings. To account for
46 multiple testing, we set the thresholds for statistical significance at $P < 5 \times 10^{-3}$.

47 **Participants:** The number of cases and controls identified in the relevant GWASs were 437 and
48 3,325 for Behçet's disease, 4,918 and 5,684 for coeliac disease, 435 and 341,188 for dermatitis
49 herpetiformis, 4,576 and 8,039 for lupus, 11,988 and 275,335 for psoriasis, 22,350 and 74,823
50 for rheumatoid arthritis, 3,597 and 337,121 for sarcoidosis, 2,735 and 332,115 for Sjögren's

51 syndrome, 9,095 and 17,584 for systemic sclerosis, 18,942 and 501,638 for type 1 diabetes,
52 2,400 and 410,350 for NHL; and 296 to 2,340 cases and 271,463 controls for NHL subtypes.

53 **Exposures:** Genetic variants predicting ten autoimmune diseases: Behçet's disease, coeliac
54 disease, dermatitis herpetiformis, lupus, psoriasis, rheumatoid arthritis, sarcoidosis, Sjögren's
55 syndrome, systemic sclerosis, and type 1 diabetes.

56 **Main outcome measures:** Estimated associations between genetically predicted susceptibility to
57 ten autoimmune diseases and the risk of NHL.

58 **Results** The variance of each autoimmune disease explained by the SNPs ranged from 0.3% to
59 3.1%. Negative associations between type 1 diabetes and sarcoidosis and the risk of NHL were
60 observed (odds ratio [OR] 0.95, 95% confidence interval [CI]: 0.92 to 0.98, $P = 5 \times 10^{-3}$, and OR
61 0.92, 95% CI: 0.85 to 0.99, $P = 2.8 \times 10^{-2}$, respectively). These findings were supported by the
62 sensitivity analyses accounting for potential pleiotropy and weak instrument bias. No significant
63 associations were found between the other eight autoimmune diseases and NHL risk. Of the
64 NHL subtypes, type 1 diabetes was most strongly associated with follicular lymphoma (OR 0.91,
65 95% CI: 0.86 to 0.96, $P = 1 \times 10^{-3}$), while sarcoidosis was most strongly associated with other and
66 unspecified NHL (OR 0.86, 95% CI: 0.75 to 0.97, $P = 1.8 \times 10^{-2}$).

67 **Conclusions** These findings suggest that genetically predicted susceptibility to type 1 diabetes,
68 and to some extent sarcoidosis, might reduce the risk of NHL. However, future studies with
69 different datasets, approaches, and populations are warranted to further examine the potential
70 associations between these autoimmune diseases and the risk of NHL.

71 **Keywords:** Autoimmune diseases; Mendelian randomisation; Non-Hodgkin lymphoma; Type 1
72 diabetes; Sarcoidosis.

73 **WHAT IS ALREADY KNOWN ON THIS TOPIC**

- 74 1. The etiology of non-Hodgkin lymphoma, a common hematological malignancy, is not
75 fully understood.
- 76 2. Observational studies have reported statistically significant associations between ten
77 autoimmune diseases (Behçet's disease, coeliac disease, dermatitis herpetiformis, lupus,
78 psoriasis, rheumatoid arthritis, sarcoidosis, Sjögren's syndrome, systemic sclerosis, and
79 type 1 diabetes) and risk of non-Hodgkin lymphoma, but these studies may be susceptible
80 to residual confounding and reverse causation.

81 **WHAT THIS STUDY ADDS**

- 82 1. Genetically predicted susceptibility to type 1 diabetes, and to some extent sarcoidosis,
83 may be associated with a reduced risk of non-Hodgkin lymphoma, while no clear
84 associations were observed between the other eight autoimmune diseases and risk of non-
85 Hodgkin lymphoma or its subtypes.

86 **HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE, OR POLICY**

- 87 1. Using an approach that seeks to address residual confounding and reverse causation,
88 these findings contradict previously reported associations between autoimmune diseases
89 and risk of non-Hodgkin lymphoma from traditional observational studies.
- 90 2. Future studies with different datasets, approaches, and populations are warranted to
91 further examine the potential associations between these autoimmune diseases and the
92 risk of NHL.

93

94

95

97 **INTRODUCTION**

98 Non-Hodgkin lymphoma (NHL), a hematological malignancy that arises from lymphocytes, is
99 one of the most common cancers, with more than 544,000 new cases every year worldwide.^{1,2}
100 Despite substantial efforts to identify risk factors for NHL, the exact etiology remains elusive.³
101 According to a recent umbrella review (i.e., systematic review of meta-analyses) evaluating the
102 associations between 134 unique environmental risk factors and the risk of NHL, ten
103 autoimmune diseases - Behçet's disease, coeliac disease, dermatitis herpetiformis, psoriasis,
104 rheumatoid arthritis, sarcoidosis, systemic lupus erythematosus (SLE), Sjögren's syndrome,
105 systemic sclerosis, and type 1 diabetes (T1D) - were identified to be statistically significantly
106 associated with an increased risk of NHL.⁴ Of these, coeliac disease, rheumatoid arthritis,
107 Sjögren's syndrome, and SLE, were classified as presenting highly suggestive (i.e., $P < 1 \times 10^{-6}$,
108 at least 1000 NHL cases, and largest study in the review reporting a nominally significant result)
109 or convincing evidence (i.e., $P < 1 \times 10^{-6}$, at least 1000 NHL cases, largest study in the review
110 reporting a nominally significant result, minimal between-study heterogeneity, and no evidence
111 of publication bias) evidence for NHL risk.^{4,5} Autoimmune diseases have long been considered
112 potential risk factors for NHL.^{6,7} Proposed mechanisms for the associations between autoimmune
113 diseases and NHL include chronic inflammation, antigen stimulation, and overlapping genetic
114 susceptibility.⁸⁻¹¹ However, the associations identified by the umbrella review were from
115 systematic reviews and meta-analyses with various study design and reporting limitations.^{4,5}
116 Furthermore, the systematic reviews and meta-analyses included only case-control and cohort
117 studies, which are susceptible to multiple biases that limit their ability to evaluate causal
118 relationships. For instance, confounding factors (e.g., socioeconomic status, family history of
119 lymphoma, and infectious diseases) and reverse causation were often not considered by the

120 previous studies.¹²⁻¹⁸ For some autoimmune diseases, such as sarcoidosis, it is often unclear if the
121 autoimmune disease precedes or develops after NHL.¹⁹⁻²⁴

122 One way of addressing the issue of residual confounding and reverse causation is through
123 instrumental variable analyses with genetic instruments, often called Mendelian randomisation
124 (MR) analyses.²⁵ Because genetic variants are randomly assigned at conception and are not
125 affected by external factors such as chronic diseases and lifestyle factors, MR analyses mimic
126 randomized experiments and are less susceptible to confounding and reverse causation compared
127 with conventional observational studies.^{26,27}

128 Therefore, the aim of this study was to use the MR design to evaluate the associations between
129 genetically predicted susceptibility to ten autoimmune diseases and the risk of NHL and NHL
130 subtypes (follicular lymphoma, mature T/NK-cell lymphomas, non-follicular lymphoma, and
131 other and unspecified types of NHL).⁴ Because nearly all observational studies on the
132 associations between autoimmune diseases and NHL suggested a positive association,^{4,9,12,13} our
133 hypothesis was that the genetically predicted susceptibility to each autoimmune disease was
134 associated with an increased risk of NHL and at least one of the NHL subtypes.

135

136 **METHODS**

137 This study is reported following the Strengthening the Reporting of Observational Studies in
138 Epidemiology Using Mendelian Randomisation guidelines (STROBE-MR, **Supplement 1**).²⁸

139 Although there is no pre-registered protocol for this study, the analyses were designed prior to
140 the conduct of the study. Data were retrieved between June 2022 and June 2023, and analyses
141 were conducted between June 2022 and October 2023. The manuscript was posted on *medRxiv*.

142

143 **Study design**

144 **Figure 1** shows a schematic summary of this two-sample MR study. Overall, we extracted
145 summary level data from genome-wide association studies (GWASs) to find genetic instruments
146 for our exposures of interest (i.e., exposure GWASs) and then investigated the associations of
147 these genetic instruments with the outcomes of interest (i.e., outcome GWASs).²⁵ For the genetic
148 instruments to be valid, three core assumptions must be met:²⁹ (1) the instruments are associated
149 with the exposure of interest (the relevance assumption), (2) the instruments are not associated
150 with any confounders of the exposure-outcome relationship (the independence assumption), and
151 (3) the instruments are associated with the outcome only through the exposure (the exclusion
152 restriction assumption).

153 The analyses were carried out in three steps: first, genetic instruments for each exposure of
154 interest (i.e., autoimmune disease) were selected from relevant exposure GWASs. Second, for
155 these genetic instruments, the genetic instrument-outcome associations were extracted from the
156 relevant outcome GWASs. For each autoimmune disease, the exposure-outcome association (i.e.,
157 the Wald ratio) was estimated by dividing the genetic instrument-outcome association by the

158 genetic instrument-exposure association. Finally, the Wald ratios were summarised using
159 different techniques (See **Statistical analyses**). Due to data availability and in order to minimize
160 confounding due to population stratification, we only considered subjects of European
161 ancestry.²⁵

162

163 **Instrumental variable selection for autoimmune diseases**

164 To fulfill the relevance assumption, we used single nucleotide polymorphisms (SNPs) as genetic
165 instruments and selected those from GWASs that were (1) strongly associated with the specific
166 autoimmune disease at genome-wide significance (i.e., $P < 5 \times 10^{-8}$) and (2) independent of one-
167 another (i.e., an $R^2 < 1 \times 10^{-3}$). We used a less stringent criterion for significance level, $P < 5 \times 10^{-6}$
168 for dermatitis herpetiformis and Sjögren's syndrome to ensure that we had at least five SNPs as
169 genetic instruments for each autoimmune disease.

170 We evaluated the ten autoimmune diseases that were at least nominally significantly associated
171 (i.e., $P < 5 \times 10^{-2}$) with risk of NHL according to a recent umbrella review:⁴ Behçet disease,
172 coeliac disease, dermatitis herpetiformis, psoriasis, rheumatoid arthritis, sarcoidosis, Sjögren's
173 syndrome, SLE, systemic sclerosis, and T1D. These autoimmune diseases were all associated
174 with an increased risk of NHL. Of these, coeliac disease, rheumatoid arthritis, Sjögren's
175 syndrome, and SLE, were classified as having highly suggestive (i.e., $P < 1 \times 10^{-6}$, at least 1000
176 NHL cases, and largest study in the review reporting a nominally significant result) or
177 convincing evidence (i.e., $P < 1 \times 10^{-6}$, at least 1000 NHL cases, largest study in the review
178 reporting a nominally significant result, minimal between-study heterogeneity, and no evidence
179 of publication bias) in the umbrella review.⁴ For each autoimmune disease, we selected the

180 GWAS that had the largest sample size. In situations where no distinct GWAS publication for an
181 autoimmune disease was available, we used summary statistics from genome-wide association
182 analyses from FinnGen Release 8 (<https://r8.finngen.fi/>, **Supplementary 2**).³⁰ Study
183 characteristics of all summary level GWAS datasets for the autoimmune diseases used in main
184 analyses are shown in **Table 1**.

185 Because human leukocyte antigen (HLA) region is highly predictive of both autoimmune
186 diseases and NHL, we excluded SNPs from this locus from all analyses in order to prevent
187 potential bias due to genetic pleiotropy.^{31,32} We defined the HLA region as location from
188 28,477,797 to 33,448,354 on chromosome 6 for Genome Reference Consortium Human Build
189 37.³³

190

191 **Outcomes**

192 Our primary outcome was risk of NHL. We selected the largest GWAS on NHL, which
193 evaluated subjects from UK Biobank and Kaiser Permanente (**Table 2** and **Supplementary**
194 **Text**).³⁴ In secondary analyses for the genetically-predicted autoimmune diseases found to be
195 nominally significantly associated with NHL (see **Statistical analyses**), we further evaluated the
196 associations between these autoimmune diseases and NHL subtypes. We identified the NHL
197 subtypes from FinnGen Release 8 and we selected four major NHL subtypes, consistent with the
198 FinnGen classification system: follicular lymphoma, mature T/natural killer (NK)-cell
199 lymphomas, non-follicular lymphoma, and other and unspecified types of NHL (**Table 2**).³⁰ Of
200 note, as the data on the NHL subtypes were extracted from a different dataset from the NHL

201 GWAS, the secondary analyses also served as replication analyses using an independent
202 dataset.³⁰

203

204 **Statistical analyses**

205 *Main analyses and sensitivity analyses*

206 For each autoimmune disease, we calculated the SNP-specific Wald ratio, defined as $\beta_{\text{EXP-OUT}} =$
207 $\beta_{\text{SNP-OUT}} / \beta_{\text{SNP-EXP}}$ (**Figure 1**). We used inverse-variance weighted (IVW) analysis as our main
208 analyses to sum the Wald ratios, which assign weights to each SNP in inverse proportion to the
209 variance of the $\beta_{\text{SNP-OUT}}$, assuming all instruments to be valid.^{35,36} However, the IVW analysis
210 may be biased if any of the included instruments are invalid (e.g., if the genetic instruments
211 affect multiple traits, which is known as horizontal pleiotropy).³⁷ Therefore, we carried out three
212 sensitivity analyses that provide unbiased estimates even in the presence of some invalid
213 instruments:^{25,38} MR-Egger regression,^{36,39} weighted mode estimator analysis,⁴⁰ and weighted
214 median estimator analysis (**Supplementary Text**).⁴¹ Furthermore, to address potential weak
215 instrument bias, which may be introduced when the genetic variants explain a very small
216 proportion of the variation in the exposure,²⁹ we included robust adjusted profile scores (RAPS)
217 as an additional sensitivity analysis (**Supplementary Text**).^{25,42}

218 We evaluated the strength of the instruments using R^2 (i.e., proportion of variance of the
219 exposure explained by the genetic instrument) and F statistics. In particular, we used the
220 `get_r_from_bsen` function in the *TwoSampleMR* package in *R* (version 4.3.0), and summed the
221 absolute values across the independent SNPs to estimate the composite R^2 for each autoimmune

222 disease. F statistics were calculated using the formula F statistic = $\left(\frac{\hat{\beta}_X}{se(\hat{\beta}_X)}\right)^2$,⁴³ where $\hat{\beta}_X$ refers
223 to the genetic association between the instrument X with the exposure, and $se(\hat{\beta}_X)$ refers to the
224 standard error of $\hat{\beta}_X$.

225 *Bidirectional analyses*

226 To help establish the direction of effects between two traits,²⁵ we carried out bidirectional
227 analyses of genetically predicted NHL and risk of each autoimmune disease.²⁵ Evidence of an
228 effect in both directions could suggest that an effect acts in both directions between two traits so
229 that changing one will change the other (i.e., the true bidirectional relationship). Otherwise it
230 may suggest that a biasing pathway may be present.²⁵ Evidence of an effect in one direction but
231 not the other supports that a biasing pathway is less likely to be present.

232 *Secondary analyses*

233 To further validate the principal findings from the main analyses, we carried out two sets of
234 secondary analyses for the autoimmune diseases that were nominally significantly associated
235 with risk of NHL (i.e., T1D and sarcoidosis, see **Results**): (1) evaluating the risk of four NHL
236 subtypes using a dataset independent from the main analyses and (2) restricting the genetic
237 instruments to those that are less likely to affect immune function. The analyses of the NHL
238 subtypes were carried out to identify which of the NHL subtypes were driving the observed
239 association with NHL, and also to serve as a replication using a separate outcome study
240 population not overlapping with the NHL study population of the other analyses. The analysis
241 involving selected pathways were carried out to address horizontal pleiotropy. In particular,
242 given the shared role of immune function in the pathway to developing both autoimmune
243 diseases and NHL, we sought to identify biological pathways for the autoimmune diseases that

244 were to a lesser degree linked to immune function (**Supplementary Table 1**). For T1D, we
245 restricted this secondary analysis to SNPs linked to insulin production, while for sarcoidosis we
246 selected SNPs that were not directly linked to immune function. The KEGG and GeneCards
247 databases were used to identify the function of all T1D and sarcoidosis SNPs.^{44,45} For all
248 secondary analyses, only IVW analyses were carried out.

249 *Software*

250 We used *TwoSampleMR* package in *R* (version 4.3.0) to run the two-sample MR analyses, and
251 *forplot* to generate the forest plots.⁴⁶ A $P < 5 \times 10^{-2}$ was considered nominally significant. To
252 correct for multiple testing of the ten autoimmune diseases in the main analyses, the level for
253 statistical significance was set at $P < 5 \times 10^{-2} / 10 = 5 \times 10^{-3}$. To make the results more interpretable,
254 all causal estimates were multiplied by 0.693 ($= \log_e^2$) and next exponentiated in order to
255 represent the odds ratios (ORs) for NHL per doubling in the prevalence of the autoimmune
256 disease under study.⁴³

257 *Ethics*

258 Only summary-level data from published studies with relevant ethical approvals were used in
259 this study so approval from institutional review board was not necessary.

260 *Patient and public involvement*

261 No patients or members of the public were involved in the conception of the study, interpretation
262 of the results, or drafting of the manuscript. We do not have plans to disseminate the results to
263 research participants or relevant patient communities.

264

265 RESULTS

266 The number of cases and controls identified in the relevant GWASs were 437 and 3,325 for
267 Behçet's disease, 4,918 and 5,684 for coeliac disease, 435 and 341,188 for dermatitis
268 herpetiformis, 4,576 and 8,039 for lupus, 11,988 and 275,335 for psoriasis, 22,350 and 74,823
269 for rheumatoid arthritis, 3,597 and 337,121 for sarcoidosis, 2,735 and 332,115 for Sjögren's
270 syndrome, 9,095 and 17,584 for systemic sclerosis, 18,942 and 501,638 for type 1 diabetes, and
271 2,400 and 410,350 for NHL (**Table 1** and **Table 2**).³⁴ The variance in the exposure explained by
272 the genetic instruments ranged from 0.3% for Behçet disease to 3.1% for T1D (**Table 1**).

273

274 Primary analyses

275 A doubling in the genetically-predicted prevalence of T1D was associated with an OR for NHL
276 of 0.95 (95% confidence interval [CI]: 0.92 to 0.98, $P = 5 \times 10^{-3}$), while a doubling in the
277 genetically-predicted prevalence of sarcoidosis was associated with an OR for NHL of 0.92 (95%
278 CI: 0.85 to 0.99, $P = 2.8 \times 10^{-2}$) (**Figure 2**). We did not observe significant associations between
279 the other eight autoimmune diseases and risk of NHL (**Figure 2**).

280

281 Sensitivity analyses

282 MR-Egger, weighted mode and weighted median yielded ORs comparable to those from the
283 main analyses, and with overlapping CIs, indicating little presence of pleiotropy (**Figure 2**).
284 Furthermore, the RAPS sensitivity analyses did not suggest bias due to weak instruments. For the
285 bidirectional analyses we did not observe significant associations between NHL and the risk of
286 any of the autoimmune diseases (**Supplementary Table 2**).

287

288 **Secondary analyses**

289 In the secondary analyses, we evaluated the autoimmune diseases that were at least nominally
290 significantly associated with risk of NHL, i.e., T1D and sarcoidosis. In the analyses restricted to
291 non-immune pathways, we observed ORs of 0.96 (95% CI: 0.87 to 1.07) for T1D and 0.96 (95%
292 CI: 0.88 to 1.06) for sarcoidosis, respectively, supporting the main analyses.

293 **Figure 3** shows the IVW analyses of genetically predicted susceptibility to T1D and sarcoidosis,
294 and the risk of NHL subtypes. For T1D, the association with composite NHL appeared to be
295 driven by follicular lymphoma, with an OR of 0.91 (95% CI: 0.86 to 0.96, $P = 1 \times 10^{-3}$).

296 Sarcoidosis was most strongly associated with other and unspecified types of NHL, with an OR
297 of 0.86 (95% CI: 0.75 to 0.97, $P = 1.8 \times 10^{-2}$).

298

299 **DISCUSSION**

300 **Principal findings**

301 In this MR study of ten autoimmune diseases previously linked to an increased risk of NHL, we
302 found that genetically predicted susceptibility to T1D, and to some extent, genetically predicted
303 susceptibility to sarcoidosis, were associated with a reduced risk of NHL. While these findings
304 were consistent across a wide range of sensitivity analyses, no clear associations were observed
305 between the other eight autoimmune diseases and risk of NHL. Using an approach that attempts
306 to address potential residual confounding and reverse causation, our findings contradict those
307 reported in previous traditional observational studies. This highlights the need for future studies
308 with different datasets, approaches, and populations to further examine the potential associations
309 between these autoimmune diseases and the risk of NHL.

310

311 **Context of primary findings**

312 Autoimmune diseases have long been considered potential risk factors for NHL, especially
313 rheumatoid arthritis, SLE, and Sjögren's syndrome.^{6,7,13,47,48} A recent large-scale prospective
314 cohort study found that among all cancers, lymphoma demonstrated the most extensive
315 associations with different immune-mediated diseases.⁹ According to a previous umbrella review
316 evaluating the associations between any environmental risk factors and the risk of NHL reported
317 in published meta-analyses, there were consistent statistically significant associations between
318 autoimmune diseases and NHL risk.⁴ Although the exact mechanisms for positive associations
319 between autoimmune diseases and NHL remain unclear, there are several mechanisms proposed,
320 including chronic inflammation, antigen stimulation, overlapping genetic susceptibility, and

321 dysfunction of certain protein families.^{8-11,49} Furthermore, the observed associations between
322 autoimmune diseases and NHL may be attributable to the immunosuppressants that are used as
323 treatments as well.^{47,50}

324 In our study, two autoimmune diseases -T1D and sarcoidosis - were found to be associated with
325 NHL. However, contrary to our pre-specified hypotheses, which were based on the summary
326 relative risks from an umbrella review (relative risk 1.55 (95% CI: 1.15 to 2.08) for associations
327 between T1D and NHL, and relative risk 1.43 (95% CI: 1.03 to 1.99) for sarcoidosis and NHL),⁴
328 we found a statistically significant negative association between genetically predicted
329 susceptibility to T1D and NHL. Although we also observed a negative association between
330 genetically predicted susceptibility to sarcoidosis and NHL, the association was no longer
331 significant after accounting for multiple testing. While no previous epidemiologic studies have
332 reported statistically significant negative associations for either T1D or sarcoidosis with
333 NHL,^{8,9,13,22,51} it has been suggested that the pathways involved in autoimmune disease and
334 cancer development may work in opposite directions.⁵² Moreover, it remains unclear if
335 sarcoidosis precedes or follows NHL, as several observational studies reported occurrence of
336 sarcoidosis among NHL patients after treatment of NHL or concomitant occurrence of two
337 diseases.¹⁹⁻²²

338 There are several potential explanations for the discrepancy between our findings and those of
339 the previous umbrella review. First, meta-analyses of observational studies and MR analyses
340 have different bias structures. While traditional observational studies often suffer from
341 confounding, misclassification, and selection bias, and meta-analyses are often susceptible to
342 publication bias (i.e., the lack of publishing certain findings, often null findings), horizontal
343 pleiotropy is of particular concern for MR studies. Second, many of the traditional observational

344 studies and meta-analyses have relatively small sample sizes. In particular, the previous meta-
345 analyses for T1D only included three individual observational studies with 1155 NHL cases and
346 the meta-analysis for sarcoidosis only included seven studies with 150 sarcoidosis cases.^{4,53}
347 Furthermore, according to a formal critical appraisal tool for systematic reviews and meta-
348 analyses (i.e., AMSTAR 2), both meta-analyses were found to have at least one critical
349 weakness.^{4,5} Third, given that genetic variants are fixed at conception, the exposure measured in
350 MR analyses are typically interpreted as the lifelong exposure. In observational studies on the
351 association between autoimmune diseases and NHL, the exposure is usually measured and
352 interpreted in a defined time period from disease onset of autoimmune diseases. Finally, we
353 conducted our study in population of European ancestry, while some meta-analyses investigated
354 the associations in multiple populations other than European ancestry.^{53,54} However, the findings
355 from populations with different ancestries did not differ drastically in the umbrella review.^{53,54}

356

357 **Strengths and limitations of this study**

358 The MR design is a major strength of our study, as it reduces the influence of residual
359 confounding and reverse causation. In particular, confounding due to infectious diseases (an
360 important risk factor for NHL and for autoimmune diseases) is minimized in this study.
361 Moreover, to our knowledge, this is the first MR study that investigates the association between
362 multiple autoimmune diseases and NHL.

363 Our study had several limitations. First, this study is limited to populations of European ancestry,
364 which have the highest incidence of NHL, and it is unclear whether the findings can be
365 generalised beyond this population. However, only GWAS data for populations of European

366 ancestry are publicly available for all the autoimmune diseases, NHL, and NHL subtypes that we
367 investigated.³ Furthermore, using the same population also ensures that the exposure dataset and
368 outcome dataset are as similar as they could be. Second, NHL is a heterogenous group of
369 haematologic disorders, with more than 20 subtypes.⁵⁰ Although we used the most robust dataset
370 for NHL subtypes (i.e., FinnGen), different categorisation of NHL subtypes is possible and may
371 yield different results. Third, one autoimmune disease, Behçet disease, had suboptimal *F*
372 statistics for its genetic instruments, which may cause weak instrument bias. However, when we
373 conducted RAPS analysis to address potential bias due to weak instruments, the observed
374 findings were similar to the main analyses.^{25,42} Lastly, given the shared genetic basis of
375 autoimmune diseases and NHL,^{11,55} we were particularly cautious about potential bias due to
376 horizontal pleiotropy by immune function. To minimize the impact of this bias, we excluded
377 SNPs in the HLA region and carried out various sensitivity analyses. We also carried out
378 secondary analyses for T1D and sarcoidosis restricted to the insulin and non-immune pathways,
379 respectively. Although our sensitivity analyses did not suggest important bias due to genetic
380 pleiotropy, it is still possible that potential pleiotropy may partially affect our study findings.
381 While our study addressed potential residual confounding that may have been present in previous
382 observational studies, the MR design may be biased if the instrumental variable assumptions do
383 not hold. Thus, the association between autoimmune diseases and risk of NHL needs further
384 investigation with evidence triangulation using different datasets, populations, and approaches.

385

386 **Conclusions**

387 This MR analysis found that genetically predicted susceptibility to T1D, and to some extent
388 genetically predicted susceptibility to sarcoidosis, were associated with a lower risk of NHL.
389 Future research using different datasets, approaches, and populations is necessary to develop a
390 more comprehensive understanding of the associations between T1D and NHL, and sarcoidosis
391 and NHL.

392 **Acknowledgements**

393 **Contributors:** XS, JDW and TR originally conceived this study. XS and TR designed this study.
394 XS and TR acquired the data. XS and TR conducted the statistical analysis. XS drafted the
395 manuscript. All authors participated in the interpretation of the data. All authors and critically
396 revised the manuscript for important intellectual content. XS and TR had full access to all the
397 data in the study and take responsibility for the integrity of the data and the accuracy of the data
398 analysis. TR provided supervision. XS and TR are the guarantors. The corresponding author
399 attests that all listed authors meet authorship criteria and that no others meeting the criteria have
400 been omitted.

401 **Funding/Support:** Dr. Rogne received funding from CTSA Grant Number UL1 TR001863 from
402 the National Center for Advancing Translational Science (NCATS), a component of the NIH.
403 The contents of this manuscript are solely the responsibility of the authors and do not necessarily
404 represent the official views of NIH. The NIH had no role in the design and conduct of the study;
405 collection, management, analysis, and interpretation of the data; preparation, review, or approval
406 of the manuscript; and decision to submit the manuscript for publication.

407 **Conflict of interest:** All authors have completed the ICMJE uniform disclosure form
408 at www.icmje.org/coi_disclosure.pdf and declare: In the past 36 months, Dr. Wallach reported
409 receiving grant support from the FDA, Arnold Ventures, Johnson & Johnson through Yale
410 University, and the National Institute on Alcohol Abuse and Alcoholism of the National
411 Institutes of Health (NIH) under award 1K01AA028258; serving as a consultant for Hagens
412 Berman Sobol Shapiro LLP and Dugan Law Firm APLC; and serving as a *medRxiv* affiliate. Dr.
413 Ma received research funding from the NIH and the Frederick A. DeLuca Foundation and served
414 as a consultant for Bristol Myers Squibb.

415 **Patient consent:** Not required

416 **Ethical approval:** Not required

417 **Data sharing:** The data will be made available via a publicly accessible repository on
418 publication. The code for analyses is included as supplementary information.

419 **Transparency:** The senior author (manuscript guarantor) affirms that the manuscript is an
420 honest, accurate, and transparent account of the study being reported; that no important aspects
421 of the study have been omitted; and that any discrepancies from the study as planned (and, if
422 relevant registered) have been explained.

423 **License:** The Corresponding Author has the right to grant on behalf of all authors and does grant
424 on behalf of all authors, a worldwide license to the Publishers and its licensees in perpetuity, in
425 all forms, formats and median (whether known now or created in the future), to i) publish,
426 reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other
427 languages, create adaptations, reprints, include within collections and create summaries, extracts
428 and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the
429 Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic
430 links from the Contribution to third party material where-ever it may be located; and, vi) license
431 any third party to do any or all of the above. The default license, a CC BY NC license, is needed.

432 This is an Open Access article distributed in accordance with the Creative Commons Attribution
433 Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt,
434 build upon this work non-commercially, and license their derivative works on different terms,
435 provided the original work is properly cited and the use is non-commercial.

436 See: <http://creativecommons.org/licenses/by-nc/4.0/>.

437 **REFERENCES**

- 438 1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J. Clin.*
439 2023;73(1):17-48.
- 440 2. Mafra A, Laversanne M, Gospodarowicz M, et al. Global patterns of non-Hodgkin
441 lymphoma in 2020. *Int. J. Cancer.* 2022;151(9):1474-1481.
- 442 3. Bispo JAB, Pinheiro PS, Kobetz EK. Epidemiology and Etiology of Leukemia and
443 Lymphoma. *Cold Spring Harb. Perspect. Med.* 2019.
- 444 4. Shi X, Zhuo H, Du Y, Nyhan K, Ioannidis J, Wallach JD. Environmental risk factors for
445 non-Hodgkin's lymphoma: umbrella review and comparison of meta-analyses of
446 summary and individual participant data. *BMJ Medicine.* 2022;1(1):e000184.
- 447 5. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic
448 reviews that include randomised or non-randomised studies of healthcare interventions,
449 or both. *BMJ.* 2017;358:j4008.
- 450 6. Zhang Y, Dai Y, Zheng T, Ma S. Risk Factors of Non-Hodgkin Lymphoma. *Expert Opin.*
451 *Med. Diagn.* 2011;5(6):539-550.
- 452 7. Bowzyk Al-Naeeb A, Ajithkumar T, Behan S, Hodson DJ. Non-Hodgkin lymphoma.
453 *BMJ (Clinical research ed.).* 2018;362:k3204-k3204.
- 454 8. Masetti R, Tiri A, Tignanelli A, et al. Autoimmunity and cancer. *Autoimmun Rev.*
455 2021;20(9):102882.
- 456 9. He MM, Lo CH, Wang K, et al. Immune-Mediated Diseases Associated With Cancer
457 Risks. *JAMA Oncol.* 2022;8(2):209-219.
- 458 10. Baecklund E, Smedby KE, Sutton LA, Askling J, Rosenquist R. Lymphoma development
459 in patients with autoimmune and inflammatory disorders--what are the driving forces?
460 *Semin. Cancer Biol.* 2014;24:61-70.
- 461 11. Conde L, Bracci PM, Halperin E, Skibola CF. A search for overlapping genetic
462 susceptibility loci between non-Hodgkin lymphoma and autoimmune diseases. *Genomics.*
463 2011;98(1):9-14.
- 464 12. Ekström Smedby K, Vajdic CM, Falster M, et al. Autoimmune disorders and risk of non-
465 Hodgkin lymphoma subtypes: a pooled analysis within the InterLymph Consortium.
466 *Blood.* 2008;111(8):4029-4038.
- 467 13. Fallah M, Liu X, Ji J, Försti A, Sundquist K, Hemminki K. Autoimmune diseases
468 associated with non-Hodgkin lymphoma: a nationwide cohort study. *Ann. Oncol.*
469 2014;25(10):2025-2030.
- 470 14. Morton LM, Slager SL, Cerhan JR, et al. Etiologic heterogeneity among non-Hodgkin
471 lymphoma subtypes: the InterLymph non-Hodgkin lymphoma subtypes project. *Journal*
472 *of the National Cancer Institute Monographs.* 2014;2014(48):130-144.
- 473 15. Op de Beeck A, Eizirik DL. Viral infections in type 1 diabetes mellitus — why the β cells?
474 *Nature Reviews Endocrinology.* 2016;12(5):263-273.
- 475 16. Chabay PA, Preciado MV. EBV primary infection in childhood and its relation to B-cell
476 lymphoma development: a mini-review from a developing region. *Int. J. Cancer.*
477 2013;133(6):1286-1292.
- 478 17. Jeong SH. HBV infection as a risk factor for non-Hodgkin lymphoma. *Lancet Oncol.*
479 2010;11(9):806.
- 480 18. Getts DR, Chastain EM, Terry RL, Miller SD. Virus infection, antiviral immunity, and
481 autoimmunity. *Immunol. Rev.* 2013;255(1):197-209.

- 482 19. London J, Grados A, Fermé C, et al. Sarcoidosis occurring after lymphoma: report of 14
483 patients and review of the literature. *Medicine (Baltimore)*. 2014;93(21):e121.
- 484 20. Cho H, Yoon DH, Kim JH, et al. Occurrence of sarcoidosis after chemotherapy for non-
485 Hodgkin lymphoma. *Korean J. Intern. Med.* 2016;31(3):605-607.
- 486 21. Chalayer É, Bachy E, Occelli P, et al. Sarcoidosis and lymphoma: a comparative study.
487 *QJM*. 2015;108(11):871-878.
- 488 22. Bonifazi M, Renzoni EA, Lower EE. Sarcoidosis and malignancy: the chicken and the egg?
489 *Curr. Opin. Pulm. Med.* 2021;27(5):455-462.
- 490 23. Hauswirth AW, Skrabs C, Schützinger C, Gaiger A, Lechner K, Jäger U. Autoimmune
491 hemolytic anemias, Evans' syndromes, and pure red cell aplasia in non-Hodgkin
492 lymphomas. *Leuk. Lymphoma*. 2007;48(6):1139-1149.
- 493 24. Jardin F. Development of autoimmunity in lymphoma. *Expert Rev. Clin. Immunol.*
494 2008;4(2):247-266.
- 495 25. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nature*
496 *Reviews Methods Primers*. 2022;2(1):1-21.
- 497 26. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-
498 forgotten) practices: the design, analysis, and interpretation of Mendelian randomization
499 studies. *Am. J. Clin. Nutr.* 2016;103(4):965-978.
- 500 27. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute
501 to understanding environmental determinants of disease? *Int. J. Epidemiol.* 2003;32(1):1-
502 22.
- 503 28. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of
504 observational studies in epidemiology using mendelian randomisation (STROBE-MR):
505 explanation and elaboration. *BMJ*. 2021;375:n2233.
- 506 29. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a
507 guide, glossary, and checklist for clinicians. *BMJ*. 2018;362:k601.
- 508 30. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-
509 phenotyped isolated population. *Nature*. 2023;613(7944):508-518.
- 510 31. Matzaraki V, Kumar V, Wijmenga C, Zhernakova A. The MHC locus and genetic
511 susceptibility to autoimmune and infectious diseases. *Genome Biol.* 2017;18(1):1-21.
- 512 32. Zhong C, Cozen W, Bolanos R, Song J, Wang SS. The role of HLA variation in
513 lymphoma aetiology and survival. *J. Intern. Med.* 2019;286(2):154-180.
- 514 33. Human Genome Region MHC.
515 <https://www.ncbi.nlm.nih.gov/grc/human/regions/MHC?asm=GRCh37>. Accessed April 7,
516 2023.
- 517 34. Rashkin SR, Graff RE, Kachuri L, et al. Pan-cancer study detects genetic risk variants
518 and shared genetic basis in two large cohorts. *Nature Communications*. 2020;11(1):4423.
- 519 35. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with
520 multiple genetic variants using summarized data. *Genet. Epidemiol.* 2013;37(7):658-665.
- 521 36. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the
522 MR-Egger method. *Eur. J. Epidemiol.* 2017;32(5):377-389.
- 523 37. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in
524 Mendelian randomization studies. *Hum. Mol. Genet.* 2018;27(R2):R195-R208.
- 525 38. Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian
526 randomization investigations. *Wellcome Open Res.* 2019;4:186.

- 527 39. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments:
528 effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.*
529 2015;44(2):512-525.
- 530 40. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian
531 randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.*
532 2017;46(6):1985-1998.
- 533 41. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian
534 randomization with some invalid instruments using a weighted median estimator. *Genet.*
535 *Epidemiol.* 2016;40(4):304-314.
- 536 42. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample
537 summary-data Mendelian randomization using robust adjusted profile score. 2020.
- 538 43. Burgess S, Thompson SG. *Mendelian randomization: methods for causal inference using*
539 *genetic variants*. CRC Press; 2021.
- 540 44. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives
541 on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 2017;45(D1):D353-D361.
- 542 45. Stelzer G, Rosen N, Plaschkes I, et al. The GeneCards Suite: From Gene Data Mining to
543 Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics.* 2016;54:1 30 31-31 30
544 33.
- 545 46. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal
546 inference across the human phenome. *eLife.* 2018;7:e34408.
- 547 47. Zintzaras E, Voulgarelis M, Moutsopoulos HM. The risk of lymphoma development in
548 autoimmune diseases: a meta-analysis. *Arch. Intern. Med.* 2005;165(20):2337-2344.
- 549 48. Cuttner J, Spiera H, Troy K, Wallenstein S. Autoimmune disease is a risk factor for the
550 development of non-Hodgkin's lymphoma. *The Journal of Rheumatology.*
551 2005;32(10):1884-1887.
- 552 49. Khanmohammadi S, Shabani M, Tabary M, Rayzan E, Rezaei N. Lymphoma in the
553 setting of autoimmune diseases: A review of association and mechanisms. *Crit. Rev.*
554 *Oncol. Hematol.* 2020;150:102945.
- 555 50. Armitage JO, Gascoyne RD, Lunning MA, Cavalli F. Non-Hodgkin lymphoma. *Lancet.*
556 2017;390(10091):298-310.
- 557 51. Mellemkjaer L, Pfeiffer RM, Engels EA, et al. Autoimmune disease in individuals and
558 close family members and susceptibility to non-Hodgkin's lymphoma. *Arthritis Rheum.*
559 2008;58(3):657-666.
- 560 52. Sakowska J, Arcimowicz Ł, Jankowiak M, et al. Autoimmunity and Cancer—Two Sides
561 of the Same Coin. *Front. Immunol.* 2022;13:793234.
- 562 53. Wang Y, Liu X, Yan P, Bi Y, Liu Y, Zhang ZJ. Association between type 1 and type 2
563 diabetes and risk of non-Hodgkin's lymphoma: A meta-analysis of cohort studies.
564 *Diabetes Metab.* 2020;46(1):8-19.
- 565 54. Simon TA, Thompson A, Gandhi KK, Hochberg MC, Suissa S. Incidence of malignancy
566 in adult patients with rheumatoid arthritis: a meta-analysis. *Arthritis Res. Ther.*
567 2015;17(1):212.
- 568 55. Din L, Sheikh M, Kosaraju N, et al. Genetic overlap between autoimmune diseases and
569 non-Hodgkin lymphoma subtypes. *Genet. Epidemiol.* 2019;43(7):844-863.

Table 1. GWAS datasets used for the autoimmune diseases in the in main analyses

Trait	Study	Countries	Cohorts	Number of cases	Number of controls	Number of SNPs	R ² (% of variance explained)	Median <i>F</i> statistics (range)
Behçet disease	Fernández et al 2021 ¹	West Europe, Italy, Spain	Meta-analysis of three independent cohorts	437	3,325	7	0.3	1 (0.1-7.9)
Coeliac disease	Dubois et al 2010 ²	The UK, the USA, Finland, Italy, Netherlands, Poland, Hungary, Spain	Meta-analysis of seven independent cohorts	4,918	5,684	12	1.2	53 (30-96)
Dermatitis herpetiformis	Kurki et al 2023 ³	Finland	FinnGen	435	341,188	8*	1.3	22 (20-35)
Psoriasis	Tsoi et al 2017 ⁴	North America, Sweden	Meta-analysis of eight independent cohorts	11,988	275,335	32	1.8	59 (30-310)
Rheumatoid arthritis	Ishigaki et al 2022 ⁵	The US, the UK, Canada, Spain, Sweden, the Netherlands, France	Meta-analysis of 25 independent cohorts	22,350	74,823	59	2.4	42 (30-762)
Sarcoidosis	Kurki et al 2023 ³	Finland	FinnGen Study	3,597	337,121	9	0.7	34 (32-116)
Sjögren's syndrome	Kurki et al 2023 ³	Finland	FinnGen Study	2,735	332,115	11*	0.8	23 (21-73)
Systemic lupus erythematosus	Wang et al 2021 ⁶	Spain, northern and western Europe, Italy	Meta-analysis of three independent cohorts	4,576	8,039	24	2.5	44 (30-268)
Systemic sclerosis	López-Isac et al 2019 ⁷	Spain, Germany, the Netherlands, USA, France, Spain, Italy, UK, Sweden, Norway, Australia/UK	Meta-analysis of 14 independent cohorts	9,095	17,584	21	1.3	41 (30-100)
Type I diabetes	Chiou et al 2021 ⁸	The USA, the UK, Ireland, Finland	Meta-analysis of nine independent cohorts including FinnGen Study	18,942	501,638	72	3.1	45 (30-1077)

Footnote: GWAS: genome-wide association study; ICD-10: International Classification of Diseases, Tenth Revision; NA: not available; the UK: the United Kingdoms; the USA: the United States.

* $P < 5 \times 10^{-6}$ used as criteria to select genetic instruments.

Table 2. GWAS datasets used for the non-Hodgkin lymphoma in the in main analyses

Trait	GWAS data source	Countries	Cohorts	Case definition	Control definition	Number of cases	Number of controls
Non-Hodgkin lymphoma	Rashkin et al 2020 ⁹	The UK, the USA	UK Biobank, Kaiser Permanente cohorts	ICD-O-3 codes and SEER cite recode paradigm ¹⁰	Individuals with no record of any cancer in the relevant registries	2,400	410,350
Follicular lymphoma	Kurki et al 2023 ³	Finland	FinnGen Study	ICD-10: C82	Population controls	955	271,463
Mature T/NK-cell lymphomas	Kurki et al 2023 ³	Finland	FinnGen Study	ICD-10: C84	Population controls	296	271,463
Non-follicular lymphoma	Kurki et al 2023 ³	Finland	FinnGen Study	ICD-10: C83	Population controls	2,340	271,463
Other and unspecified types of non-Hodgkin lymphoma	Kurki et al 2023 ³	Finland	FinnGen Study	ICD-10: C85	Population controls	982	271,463

Footnote: GWAS: genome-wide association study; ICD-10: International Classification of Diseases, Tenth Revision; NA: not available; UK:

United Kingdoms; the USA: the United States.

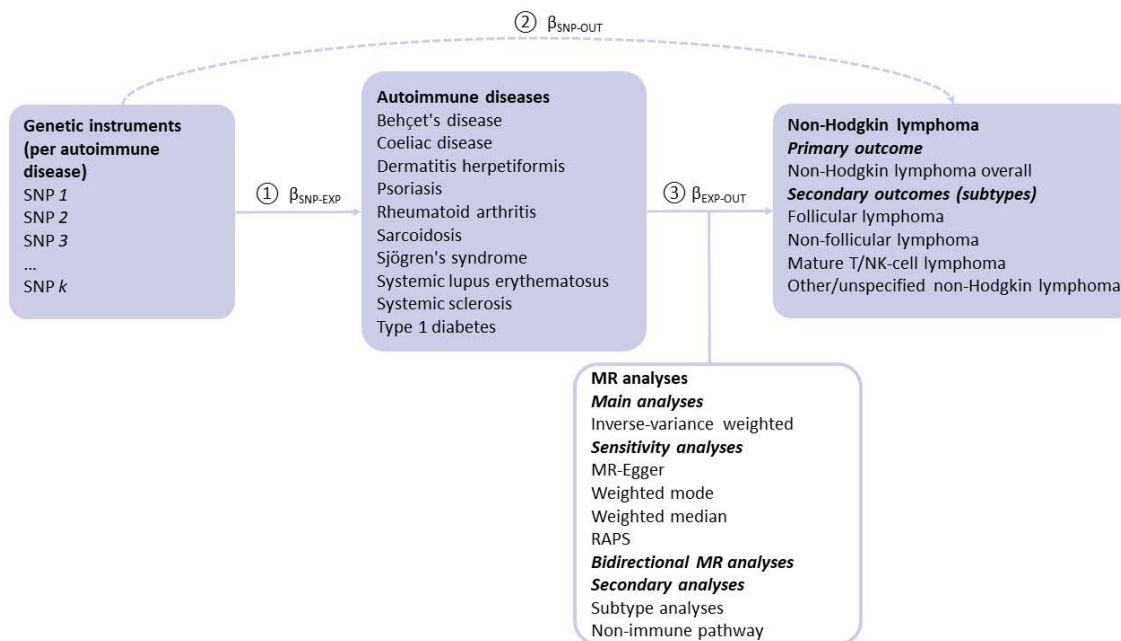


Figure 1. A schematic overview of the study design

Abbreviations: EXP: exposure; HLA: human leukocyte antigen; MR: mendelian randomisation; OUT: outcome; RAPS: robust adjusted profile scores; SNP: single nucleotide polymorphism.

Wald ratio: $\beta_{\text{EXP-OUT}} = \frac{\beta_{\text{SNP-OUT}}}{\beta_{\text{SNP-EXP}}}$; The analyses were carried out in three steps: Step 1, genetic instruments for each autoimmune disease were selected from relevant GWASs (i.e., $P < 5 \times 10^{-8}$ and an $R^2 < 0.001$), with $\beta_{\text{SNP-EXP}}$ (i.e., association between genetic instrument and the exposure) extracted. Step 2, for these genetic instruments, the genetic instrument-outcome associations $\beta_{\text{SNP-OUT}}$ were extracted from the outcome GWAS. Step 3, for each autoimmune disease, the Wald ratio (i.e., the causal estimate) was estimated for k number of SNPs by the formula $\beta_{\text{EXP-OUT}} = \beta_{\text{SNP-OUT}} / \beta_{\text{SNP-EXP}}$, and next summarised using the listed MR analyses.

In the secondary analyses, only inverse-variance weighted method was applied.

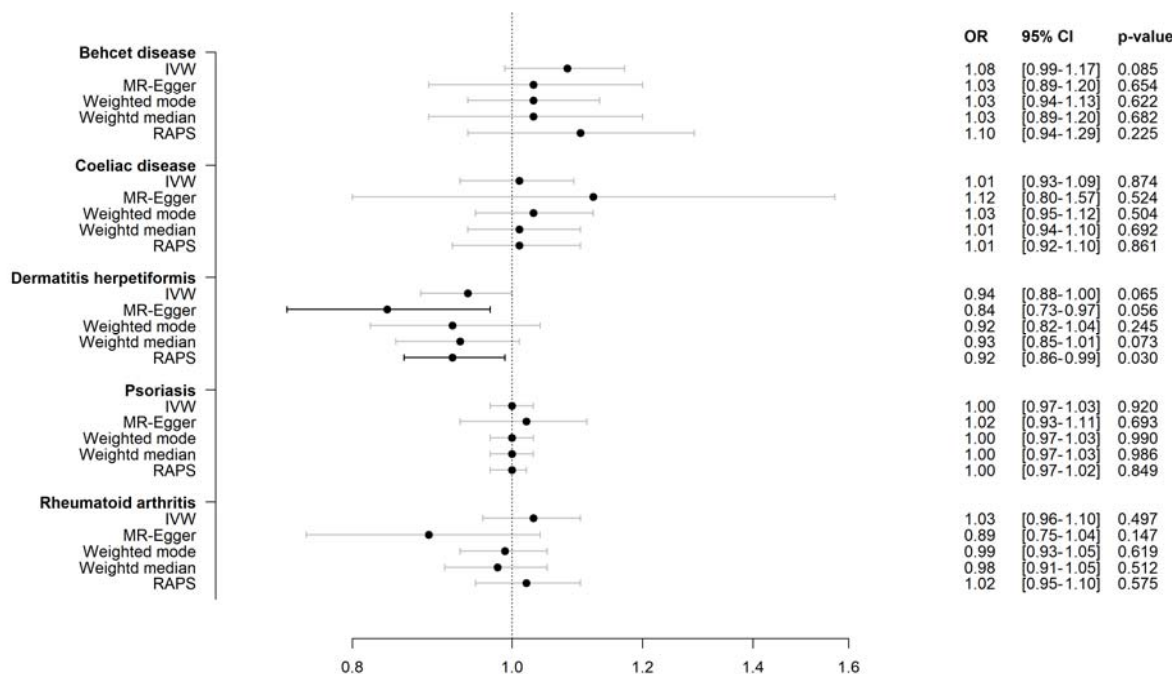


Figure 2a. Genetically predicted susceptibility to ten autoimmune diseases and risk of non-Hodgkin lymphoma overall

Abbreviations: CI: confidence interval; IVW: inverse-variance weighted; MR: mendelian randomisation; OR: odds ratio; RAPS: robust adjusted profile scores.

Legend: all causal estimates were multiplied by 0.693 and next exponentiated in order to represent the odds ratios (ORs) for NHL per doubling in the prevalence of the autoimmune disease under study.

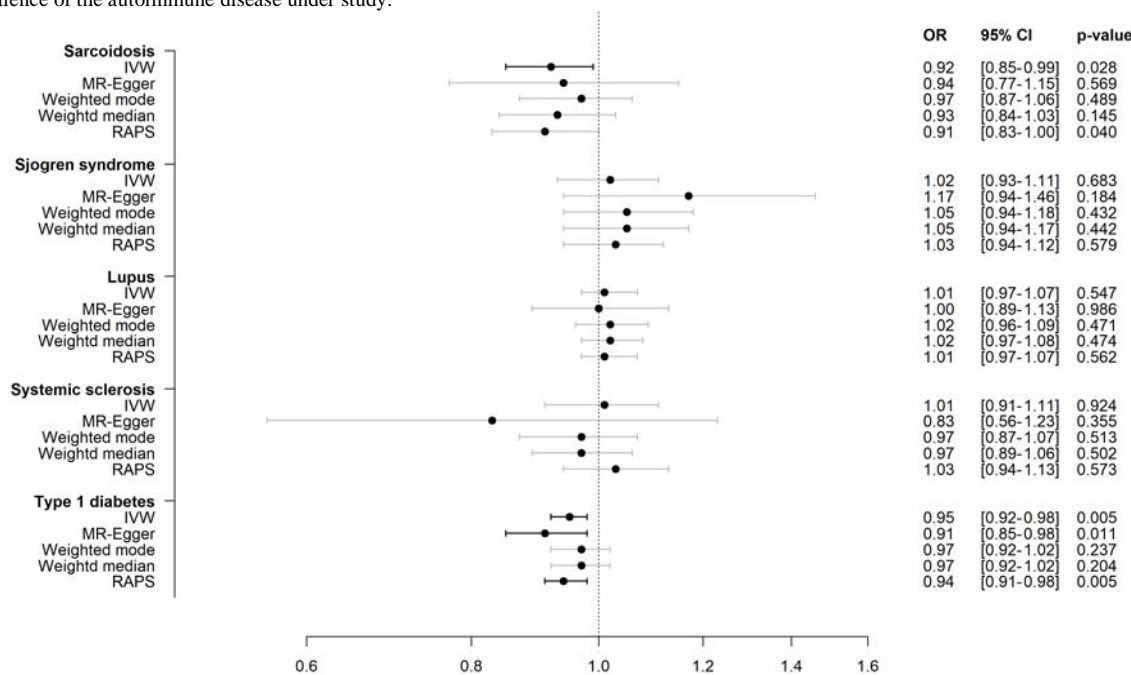


Figure 2b. Genetically predicted susceptibility to ten autoimmune diseases and risk of non-Hodgkin lymphoma overall

Abbreviations: CI: confidence interval; IVW: inverse-variance weighted; MR: mendelian randomisation; OR: odds ratio; RAPS: robust adjusted profile scores.

Legend: all causal estimates were multiplied by 0.693 and next exponentiated in order to represent the odds ratios (ORs) for NHL per doubling in the prevalence of the autoimmune disease under study.

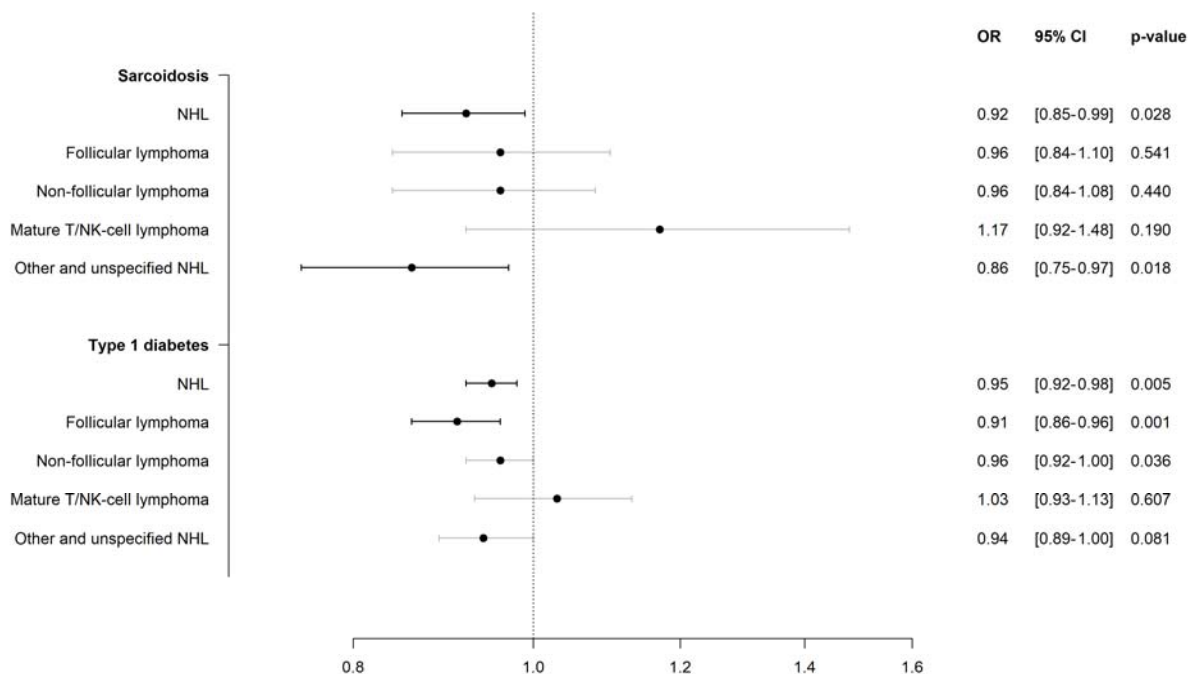


Figure 3. Inverse-variance weighted analyses of genetically predicted susceptibility to type 1 diabetes, and sarcoidosis, and risk of non-Hodgkin lymphoma subtypes

Abbreviations: CI: confidence interval; NHL: non-Hodgkin lymphoma; OR: odds ratio.

Legend: 1. the analyses were based on inverse-variance weighted approach. 2. all causal estimates were multiplied by 0.693 and next exponentiated in order to represent the odds ratios (ORs) for NHL per doubling in the prevalence of the autoimmune disease under study. 3. As mentioned in Table 2, we used different datasets for NHL and NHL subtypes.

References

1. Ortiz Fernández L, Coit P, Yilmaz V, et al. Genetic Association of a Gain-of-Function IFNGR1 Polymorphism and the Intergenic Region LNCAROD/DKK1 With Behçet's Disease. *Arthritis & Rheumatology*. 2021;73(7):1244-1252.
2. Dubois PCA, Trynka G, Franke L, et al. Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet*. 2010;42(4):295-302.
3. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613(7944):508-518.
4. Tsoi LC, Stuart PE, Tian C, et al. Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. *Nature Communications*. 2017;8(1):15382.
5. Ishigaki K, Sakaue S, Terao C, et al. Multi-ancestry genome-wide association analyses identify novel genetic mechanisms in rheumatoid arthritis. *Nat. Genet*. 2022;54(11):1640-1651.
6. Wang Y-F, Zhang Y, Lin Z, et al. Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups. *Nature Communications*. 2021;12(1):772.
7. López-Isac E, Acosta-Herrera M, Kerick M, et al. GWAS for systemic sclerosis identifies multiple risk loci and highlights fibrotic and vasculopathy pathways. *Nature Communications*. 2019;10(1):4955.
8. Chiou J, Geusz RJ, Okino M-L, et al. Interpreting type 1 diabetes risk with genetics and single-cell epigenomics. *Nature*. 2021;594(7863):398-402.
9. Rashkin SR, Graff RE, Kachuri L, et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nature Communications*. 2020;11(1):4423.
10. SEER Site Recode. https://seer.cancer.gov/siterecode/icdo3_dwhohome/index.html. Accessed April 6, 2023.