0.0.0.01em *

0pt3.25ex plus 1ex minus .2ex1.5ex plus .2ex

Coeliac Disease and Non-Hodgkins Lymphoma: A Mendelian Randomisation Study

Ahmed Abdulaal, Aimee Tham, Adam Jackson, Irene Sugurova

Translational Data Science

Department of Public Health
Imperial College London

March 26, 2021

Abstract

INSERT ABSTRACT HERE

Contents

List of Tables

List of Figures

1 Introduction

1.1 Background

Non-Hodgkin's lymphoma (NHL) represents a heterogeneous group of malignancies of the lymphoid arm of the immune system[**?**] which are associated with several recognized risk factors including male sex, age above 50 years, immunodeficiency states including post-organ transplant[**?** **?**], coeliac disease[**?**], and several viral agents such as Epstein Barr virus (EBV)[**?**] and human T-lymphocytotrophic virus-1 (HTLV-1)[**?**]. In particular, enteropathy associated t-cell lymphoma (EATL) is a rare NHL with a poor prognosis of less than one year, and a high association with coeliac disease.[**?**]

Whilst there have been previous studies to suggest an association between coeliac disease and NHL that is not of the EATL type, these studies have been observational,[**?**] and were limited to analyses of small sample sizes[**?**]. Mendelian randomisation (MR) is an analytical technique which utilises genetic information in the form of single nucleotide polymorphisms (SNPs) as an instrumental variable to investigate the link between modifiable risk factors and disease-states.[**?**] This is significant because it can surmount the limitation of unmeasured confounding in observational analyses[**?**], and therefore support causal inferences about the effects of the risk factor under investigation[**?**]. We utilised a two-sample MR approach to investigate the possible causal link between coeliac disease and NHL. To our knowledge, no such MR analysis has previously been performed to assess the effect of coeliac disease on NHL.

1.2 Objectives

The main objective of this study was to use Mendelian Randomisation (MR) to assess a causal relationship between Coeliac disease and NHL. Two techniques were used in this study to improve the robustness of the results and to take advantage of the power of the results in [**?**].

2 Methods

2.1 Genetic data on coeliac disease

Thirty-five SNPs were associated with coeliac disease at genome-wide significance of P<5e-8 in a cohort of 11,489 seropositive coeliac cases and 22,308 healthy controls in a publicly available genome-wide association study (GWAS) by Marquez et al.[**?**] The study was conducted in 2018 and the coeliac disease cohort consisted of European participants from Spain, Italy, the United Kingdom (UK), and the Netherlands. Controls were matched on ethnicity. Genotyping was performed with Immunochip (Illumina, CA), a custom dense genotyping array. Data was imputed using the 1000 genomes Phase III as a reference panel.[**?**] SNPs were excluded if they were not in Hardy-Weinberg equilibrium (P<0.001), had low minor allele frequency (<0.01) or a low call-rate (<98%) both prior to, and following, genome imputation. Relatedness in the data was mitigated by calculating the first 10 principle components (PCs) of markers associated with ancestry on the immunochip and including these as covariates when performing logistic regression.

2.2 Genetic data on NHL

Following clumping of the identified SNPs at $r^2 < 0.01$ using a 10 mega-base (mb) window, we extracted 23 of the significant SNPs from the UK Biobank (UKB). UKB is a cohort study of XXXXXXX participants aged between 38 to 73 years, enrolled between 2006 to 2010.[**?**] NHL cases were identified from international classification of disease (ICD10 and ICD9) codes, which are accrued through linkage to the National Cancer Registry at 31 discrete time-points (UKB data field ID: 40006).[**?**] There were XXX cases of NHL, and XXX controls. UKB participants were genotyped with the Affymetrix UK Biobank Axiom array,[**?**] and genomic imputation was performed using the 1000 genomes phase III, the Haplotype Reference Consortium, and UK10K as reference panels.[**?**] We accounted for relatedness in UKB by including the first 10 PCs associated with ancestry as covariates. Additionally, coeliac disease status (ICD10 code: K900) was accrued from linkage of UKB data with the Hospital Episode Statistics (HES) database.

2.3 Assumptions and sensitivity analyses

Valid instrumental variables (IVs) follow three main assumptions. Firstly, the relevance assumption states that IVs must associate with the risk factor of interest, in this case coeliac disease. To satisfy this assumption in the context of a two-sample MR, we note that the selected variants are associated with coeliac disease in a large genome-wide study [**?**]. Furthermore, the discriminative ability of a PRS score composed from the individual genetic variants is assessed using the area under the receiver operating characteristic curve (Figure XX). Secondly, the independence assumption states that IVs should not share common cause with the outcome. In this context, population stratification can also induce correlation between genotype and confounders. To mitigate for this, we adjust for principle components to account for relatedness within the UKB. Additionally, we note that whilst some of the SNPs identified by Marquez et al are also associated with scleroderma and rheumatoid

arthritis, these conditions do not represent classical risk factors for NHL [**?**]. Finally, the exclusion restriction assumption states that the IVs must not affect the outcome except through the risk factor. To investigate for the presence of horizontal pleiotropy, we utilise the I^2 statistic and Cochran's Q test as sensitivity analyses to quantify heterogenity in effect sizes between individual genetic instruments [**?**].

In order to further assess and mitigate for the latter two MR assumptions, we utilise MR Egger regression [**?**], weighted mode [**?**] and weighted median [**?**] analyses, and the MR pleiotropy residual sum and outlier test (MR-PRESSO) [**?**].

Assumption 1. Genetic variants are associated with the exposure.

Assumption 2. Genetic variants are independent of any confounder of exposure-outcome.

Assumption 3. Genetic variants are independent of the outcome.

2.4 Statistical methods: main analysis

2.4.1 Observational Analysis

To replicate findings in the literature [ref], we performed an association analysis of coeliac disease diagnosis on NHL diagnosis using logistic regression. Possible confounding was mitigated by including known NHL risk factors: age, sex, immunodeficiency (ICD10 codes: D80-D89) and HIV diagnosis (ICD10 codes: B20-B24) (BMJ Best Practice Ref). Inclusion of the first 10 Genetic PCs controlled for confounding by ancestry.

2.4.2 Two Sample Individual MR

Polygenic Risk Scores (PRS) were calculated for the 487,297 participants for which there was genotype data. Risk at a given loci was defined using an additive genetic model, such that risk allele homozygous = 2, heterozygous = 1, and non-risk allele homozygous = 0. Weights were the effect sizes from the Marquez et al. [**?**] GWAS and are shown in Supplementary Table . Use of external summary data necessitated a quality assurance step to match risk alleles with UK Biobank genotype data. Individual's PRS were calculated as the weighted sum of the risk alleles:

$$wAS_k = \sum_{i=1}^{n} w_i G_{k,i}$$

where $i = 1, ..., n$ index of genetic variants

$k = 1, ..., m$ number of participants

wAS_k is the weighted sum PRS for individual k

w_i is a vector of the log effect estimates

$G_{k,i}$ is a vector of the additive genetic risk for individual k

Causative effects of risk factor on outcome can be assessed using the two-stage least squares technique, which requires both exposure and outcome to be continuous. As both coeliac

disease and NHL are binary variables, we adapted this model to a two-stage logistic regression. Stage 1 regresses the participant PRS against coeliac disease diagnosis and predicts the odds of a given individual of having coeliac disease. Stage 2 regresses the predicted odds from stage 1 against NHL diagnosis. Confounding in stage 2 was addressed by including the following covariates in the model: age, sex and the first 10 genetic principle components.

Predictive performance in stage 1 was assessed using Area Under the Curve (AUC) and the F1 score. As we were only interested in a causal relationship, and not effect size, we used the p-value of the coeliac predictions on NHL to assess causation. We also assessed the residual deviance between models with and without the predicted coeliac odds.

Missing genotype data warranted a sensitivity analysis using a weighted mean PRS, such that $wAS_k = \frac{1}{n}\sum_{i=1}^{n} w_i G_{k,i}$

The biological mechanisms behind coeliac disease make it preferable for those afflicted to be gluten-free. Being gluten-free can therefore be used as a positive control to test the validity of the PRS using the same method described in stage 1. Participants were asked in 5 instances from April 2009 to June 2012 what there typical diet was in the past 24 hours (Field ID = 20086). Participants who had reported following a gluten-free diet in at least one instance were considered gluten-free.

2.4.3 Two-sample summary MR

2.4.3.1 *IVW*

2.4.4 need to define subsubsubsection - IVW

A two-sample summary MR was run using SNP data from a meta-analysis GWAS on coeliac disease [**?**] which identified 23 SNPs that are independently associated with coeliac disease. These SNPs were chosen with regards to a p-value threshold of 5 x 10-8 and a linkage disequilibrium threshold of 0.001. These SNPs were then harmonised, and a summary MR using UK Biobank data was run. Individual ratio estimates of the change in log odds of NHL per unit (allele) change in the log odds of coeliac disease is defined as:

$$\hat{\theta}_j = \frac{\beta_{Y_j}}{\beta_{X_j}}$$

Where β_{Xi} is the per allele change in the log odds of coeliac disease, and β_{Yi} is the change in log odds of NHL based on a unit change of β_{Xi}. The ratio estimates are then combined into an inverse variance weighted (IVW) estimate as follows:

$$\hat{\theta}_{IVW} = \frac{\sum_{j=1}^{n} \hat{\theta}_j / var(\hat{\theta}_j)}{\sum_{j=1}^{n} 1/var(\hat{\theta}_j)}$$

Where $var(\hat{\theta}_j)$ is the variance of the ratio estimate. The IVW therefore represents a weighted average of slope estimates, and provides a weighted estimate of the potential causal effect.

2.4.5 MR Egger

MR Egger is used to account for directional pleiotropy by estimating an intercept term. This technique allows all variants to have pleiotropic effects, provided they are not proportional to the variants' effects on the risk factor of interest. In other words, it requires that pleiotropic effects are independent of instrument strength. This is known as the instrument strength independent of direct effects (InSIDE) assumption [**?**]. Horizontal pleiotropy is indicated by a non-zero intercept [**?**].

2.4.6 Weighted mode

Mode-based estimation is grounded on the zero modal pleiotropy assumption, which is that the most common horizontal pleiotropy value across IVs is zero. This technique allows for between 50 - 100% of genetic instruments to demonstrate pleiotropic effect. Additionally, the weighted mode is robust to outliers and remains consistent if the largest weights are contributed by valid IVs [**?**].

2.4.7 Weighted median

Median-based estimation is valid when the majority of IVs are valid [**?**]. This method additionally assumes symmetric pleiotropy. When this is the case, the median of the ratio estimates represents a sensible estimate of the causal link.

2.4.8 MR PRESSO

2.5 Assessment of assumptions

2.6 Sensitivity analyses

Under two-sample summary MR, used leave one out analysis as a sensitivity test.

2.7 Software and pre-registration

Base R function glm.

3 Results

3.1 Observational Analysis

Table 1: Descriptive 1

	Controls	NHL Cases
n	461298	3394
Male (%)	204230 (44.3)	1861 (54.8)
Age (year)		
Min	49	50
Median (IQR)	69 (13)	74 (9)
Max	86	83
Coeliac (%)	2836 (0.6)	37 (1.1)
Immunodeficient (%)	1795 (0.4)	193 (5.7)
HIV (%)	189 (0.0)	9 (0.3)

Study eligibility required consent and availability of genotype data. Removal of 22,605 individuals from the control group that were diagnosed with non-NHL cancer was predicated on the hypothesis that there might be some underlying biological association with NHL (Supplementary figure 6). Analysis was conducted on 3394 NHL cases and 461298 controls that fit this criteria (Supplementary figure 6).

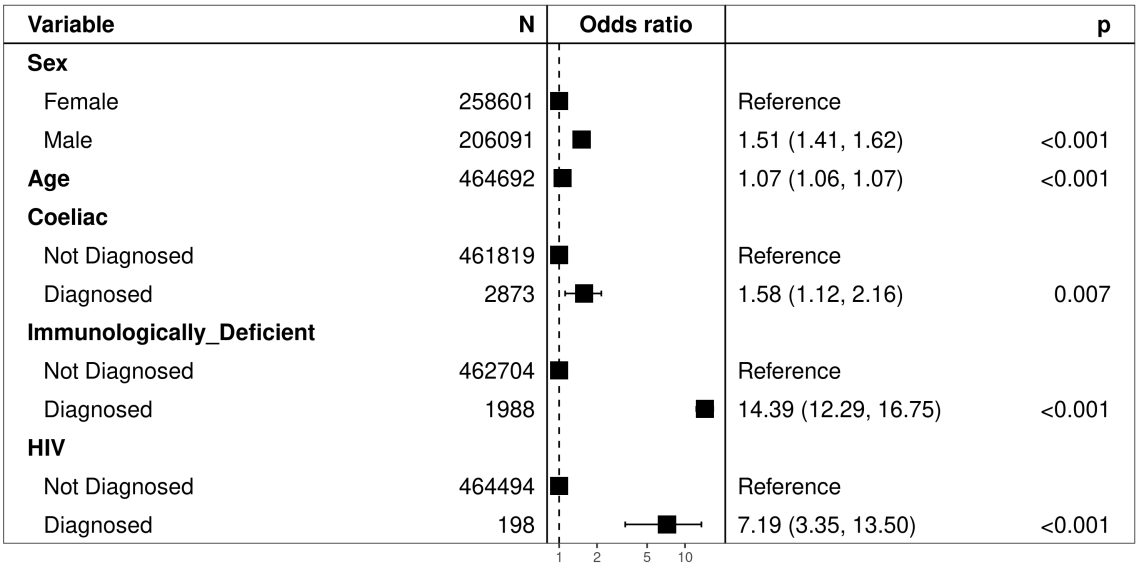
There was a greater proportion of male cases (54.8%) than controls (44.3%) (Table 1). Cases had a higher median age (median=74, IQR=9) than controls (median=69, IQR=13) (Table 1). As seen in Table 1, these is a higher proportion of our study population with coeliac disease, an immunologically deficient disease, and HIV. Previous research has shown that coeliac disease often occurs concurrently with other autoimmune diseases [] Lundin so it is important to consider whether the association is with coeliac disease or with other autoimmune disease that occur in tandem. A visual comparison of the first three genetic PCs for cases and controls suggests virtually no confounding by ancestry (Supplementary figure 2).

Results from a multivariable logistic regression on NHL Lymphoma diagnosis suggests a significant positive association between coeliac disease and NHL (OR=1.58, 95% CI=[1.12-2.16], p-value=0.007) (Figure). The association between coeliac disease and NHL was lower than for other autoimmune diseases and HIV. Analysis without the covariates immunodeficiency and HIV diagnosis suggest that these do not influence the effect of coeliac disease on NHL (OR=1.58, 95% CI=[1.12-2.16], p-value=0.007).

3.2 Two Sample Individual MR

Table highlights the distribution of the Polygenic Risk Score (PRS) across different cohort groups, and is further visualised for coeliac disease in Figure . Mean PRS was greater in those diagnosed with coeliac disease (mean=0.3163, SD = 0.5279) than those not diagnosed (mean=0.0927, SD=0.5262). Mean PRS was also greater in those who were Gluten-free (mean=0.1455, SD=0.5310) than those with No Special Diet (mean=0.0934, SD=0.5264).

Figure 1: Logistic regression of selected covariates on NHL. The first 10 genomic Principle Components were also included in the regression but are not shown here.



The PRS had some discriminative ability to predict coeliac disease (AUC=0.618, Figure). Mention PRS effect sizes and standard error. Use of a mean PRS instead of the summative PRS also produced an AUC of 0.618, suggesting that missing genotype data had no impact on model performance. The PRS was further validated using a postive control, having some power in predicting whether an individual was gluten-free (AUC=0.53).

In stage 2 of the two-stage logistic regression we regressed the predicted odds for coeliac disease against NHL, with age, sex and the first 10 genetic PCs as covariates. The predicted coeliac odds were not significant (OR=21.81, 95% CI=[0, 4590897.61], p-value=0.6). ANOVA of a model with only the covariates, and a model with the covariates and the coeliac predictions had a residual deviance of 0.2381. Together these findings suggest no causative effects of coeliac disease on NHL.

3.3 Main results

3.4 Assessment of assumptions

3.5 Sensitivity and additional analyses

Table 2: Descriptive 2

	n	Mean	SD	Min	Max
Cohort	464692	0.0941	0.5265	-2.1954	2.6841
Coeliac					
Diagnosed	2873	0.3163	0.5279	-1.5176	2.4968
Not Diagnosed	461819	0.0927	0.5262	-2.1954	2.6841
NHL					
Diagnosed	3394	0.0955	0.5294	-1.7202	2.0575
Not Diagnosed	461298	0.0941	0.5265	-2.1954	2.6841
Sex					
Female	258601	0.0940	0.5260	-2.1954	2.4136
Male	206091	0.0942	0.5270	-2.1675	2.6841
Diet					
No Special Diet	409087	0.0934	0.5264	-2.1756	2.6841
Gluten free	5628	0.1455	0.5310	-1.6806	2.1497
Lactose free	4309	0.1071	0.5109	-1.6806	2.0223
Low Calorie	32929	0.0933	0.5261	-2.0076	2.2640
Vegetarian	8553	0.1015	0.5238	-1.6867	2.1045
Vegan	841	0.0948	0.5303	-1.5535	1.6326
Other	13915	0.0973	0.5282	-2.1954	2.0641

4 Discussion

4.1 Key results

Observational analysis in this study shows an association between coeliac disease and NHL, even accounting for immuno-compromised disease, which is consistent with previous study findings. (Will look into biomarkers that may be impacting this)

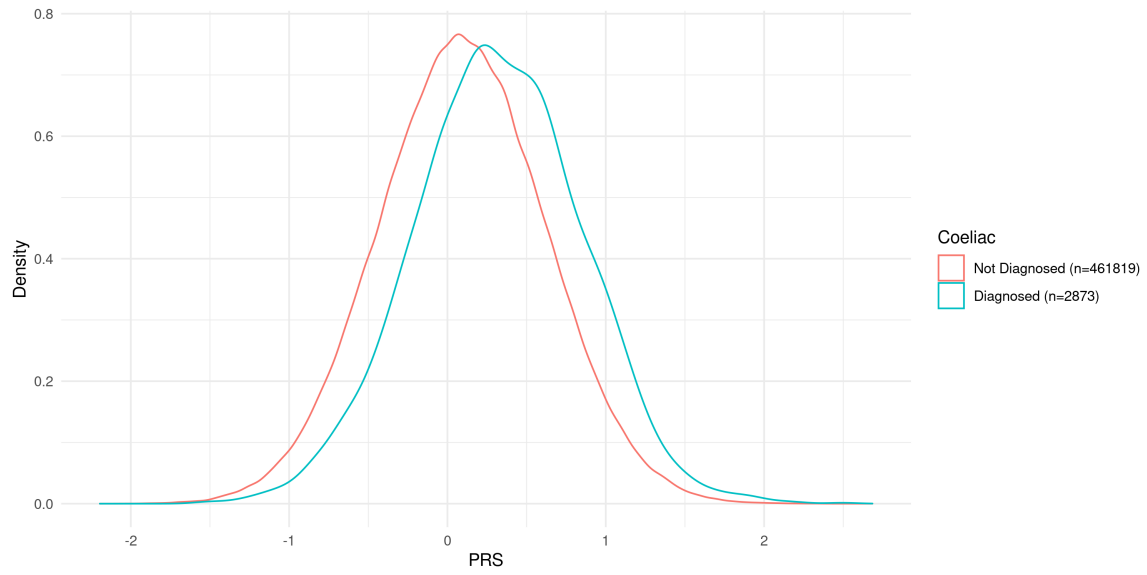
Both two-sample individual MR and two-sample summary MR techniques did not provide evidence of a causal relationship between coeliac disease and NHL, which suggests there is a unknown confounder causing this associative relationship.

4.2 Strengths

This study was the first of it's kind to investigate a causative link between coeliac disease and NHL. Use of an independent summary level dataset for genetic instruments increased the power of the study. Marquez et al. coeliac GWAS was the largest available in European populations. Use of UK Biobank is one of the largest available cohorts to investigate a wide range of diseases including NHL.

We used two MR techniques: two-sample individual MR and two-sample summary MR, and a comprehensive set of sensitivity analyses. This allowed us to explore the limitations of each model and account for violations of the MR assumptions. It's reassuring that all techniques resulted in the same outcome.

Figure 2: PRS for Coeliac diagnosis



4.3 Limitations

4.3.1 Information Bias

Coeliac disease diagnosis involves blood tests, an endoscopic biopsy and eating foods containing gluten [ref]. This high barrier to diagnosis means that many patients take years to be diagnosed and some are never diagnosed at all. This misclassification bias will undoubtedly be reflected in both the Marquez and UK Biobank cohorts, but is difficult to control for.

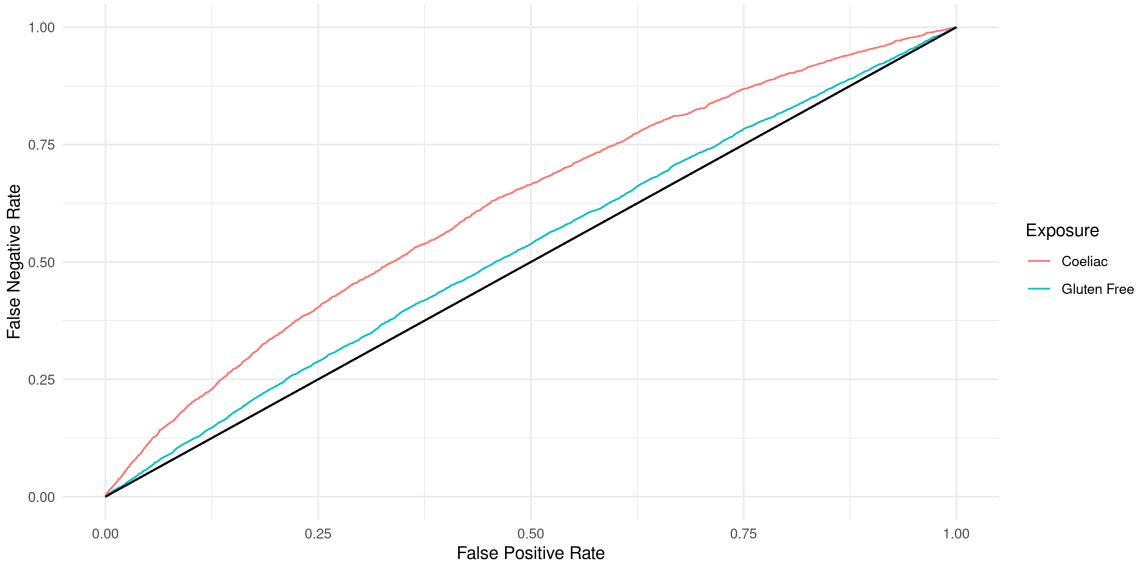
4.3.2 Two-sample individual MR Limitations

The weighted sum PRS in two-stage MR is particularly vulnerable to horizontal pleiotropy. To attenuate this, we applied a more robust clumping threshold of 0.001 than applied by Marquez et al. Novel Bayesian regression PRS approaches that apply a continuous shrinkage prior in a multivariate framework can deal with this directly (Ge et al. 2019). The MR-EGGER approach used in our second analysis allows for directional pleiotropy by introducing an intercept to the model. There is also increasing evidence to show that variants that do not meet the typical Bonferroni threshold of significance can help to predict a given trait, use of >23 coeliac-associated SNPs may be important for future studies [ref].

We adapted the two-stage least squares approach to a two-stage logistic regression. It's important to incorporate large standard errors from stage one into stage two; R-packages such as ivprobit can do this directly. However, the PRS predictors in the first logistic regression had small standard errors making use of these packages redundant.

Coeliac disease is recorded as a binary trait, but it generally presents as a spectrum of severity and symptoms, so may be better represented as a continuous score. Burgess and Labrecque. 2018 demonstrated that dichotomisation of a continuous risk factor can lead to violation of the exclusion restriction assumption; the genetic instruments can impact the outcome independently of a change in the risk factor. We mediate this to an extent by using

Figure 3: PRS ROC for Coeliac diagnosis



a continuous predicted value for coeliac disease in the two-sample individual MR technique. Characterising binary traits using continuous PRS is proving to be an increasingly popular avenue for researchers.

All individuals who reported being gluten free the previous day in 1 of 5 instances over n years was classified as gluten-free. This measure is not robust: only capturing diet from the previous day and being susceptible to short term health trends. We validated this positive control approach by confirming no predictive ability of the PRS for other special diets such as being vegetarian or lactose free.

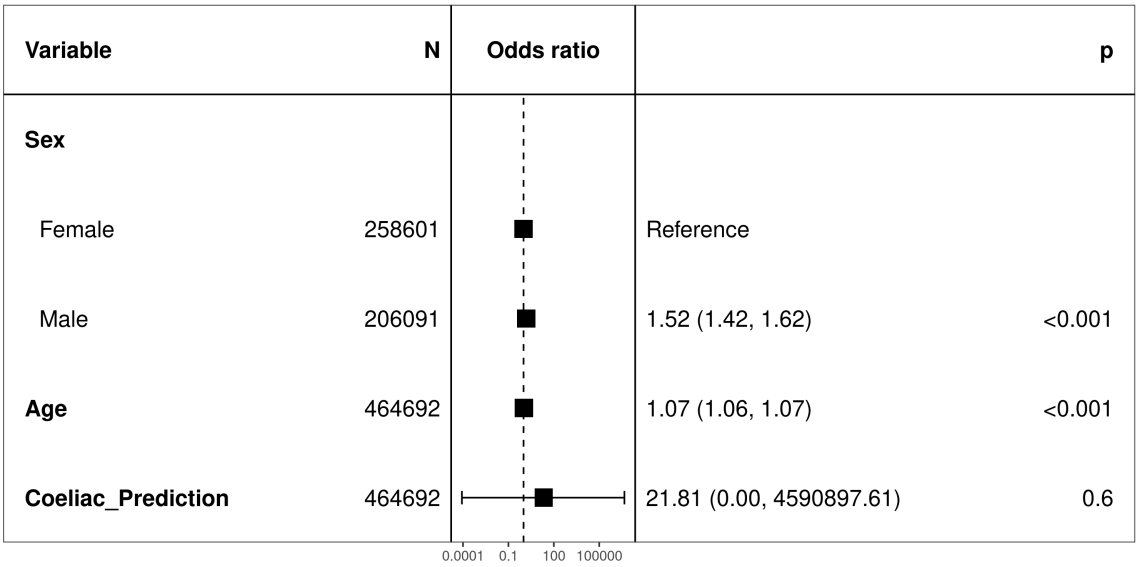
4.3.3 General MR Limitations

As Marquez et al. used a European population, and we used a UK population, it's possible that there are some immeasurable differences in population structure between the two samples.

Potential reasons for not finding an association

- Another confounder that explains the observational association between coeliac disease and NHL. Likely some sort of immune pathway
- Dichotimisation of coeliac disease when it is in fact a continuous measure
- Power issue. Unlikely due to the use of two independent large cohorts of 30k and 460k
- It's also possible that the impact of coeliac disease on NHL is related to severity and/or time-dependent damage caused by having coeliac disease.
- We assume random distribution of genetic variants BUT this could be false

Figure 4: Stage 2 Forest Plot



- social factors could influence the genes we inherit
- people do not choose baby daddies at random lol – assertive mating
- parental non-genetic characteristics can influence outcomes eg educated dad would encourage child education – dynastic effects – so could do within sibling analysis to avoid this bias
- there are geographical patterns in genetics, but no solution exists yet
- We assume the effect of changing the genetic variants is the same as the effect of changing the modifiable exposure we are interested in through other means, eg we assume that changing BMI by changing the number of BMI-increasing alleles a person has is going to have the same effects on the outcome as changing BMI through interventions we are interested in such as diet or exercise – gene-environment equivalence. The more biologically proximal the exposure is (how close the exposure is to a protein made by only a few genes), the more likely the gene- environment equivalence assumption is to hold
- Effect estimate does not necessarily map directly on to the potential effect of a clinical or public health intervention. An important reason for this is that Mendelian randomisation estimates a 'lifetime' effect of the exposure. For example, if our exposure is systolic blood pressure and our outcome is cardiovascular disease events, we can estimate the effect of having a lower systolic blood pressure by 10 mmHg across the entire life course. Whilst this can tell us something about whether reducing systolic blood pressure is likely to have an effect on the frequency of cardiovascular disease events, it cannot give us an exact answer about how much we will change the incidence of cardiovascular disease events if we administer antihypertensives to adults in mid-life

4.4 Interpretation

Clinical Relevance

- Association does not equal causation. But even if there was a causative effect how it would manifest itself is unclear. Is it simply having coeliac disease that would cause NHL? Or is it continuing a lifestyle that excacerbates the symptoms of coeliac disease that eventually results in NHL? This is an important question as it potentially suggest different treatment paths. Early diagnosis of coeliac disease and removal of gluten from the diet could potentially help prevent NHL further down the line. But if the causative effects of coeliac disease on NHL is independent of gluten intake then clinicians should focus on early diagnosis of NHL instead. Use of gluten as an instrumental variable could be a good idea. Ideally a clinical trial would be the most appropriate, but may involve some ethical considerations. Gluten intake also varies between cultures, but using culture/gluten intake as an instrumental variable could enable confounding by ancestry and violation of the exclusion restriction assumption.

4.5 Generalisability

There is a concern that the population of UK Biobank may not be representative due to the fact that only 6 percent of those invited chose to participate in the study (ref to Batty). Batty et al. (2020) investigated whether this population was representative of the entire UK population, in particular with regards to the association between risk factors and mortality. The findings of this study were that risk factor associations for a number of mortality causes were representative of the population of England and Scotland, meaning that the results of our study should be generalisable.

However, there are also concerns about the generalisability of results across ethnic groups. UK Biobank is predominantly made up of white European ancestry, and the GWAS data from [**?**] contains only those of white European ancestry, so the results found in this analysis is unlikely to be matched among those of different ethnic background. This analysis would need to be undertaken in other populations for it to be generalisable.

5 Conclusion

Group 8 is amazing. - agree

This study confirms the observational association between coeliac disease and NHL that has been established in previous studies. However, using MR methods has not provided evidence of a causal relationship. There is scope for further investigation into potential confounders of this associative relationship.

6 SUPPLEMENTARY STUFF

Make a note of what figures you want to be added to this section. For instance, figure 8a.

7 FIGURES

Figure 5: YOUR CAPTION

Figure 6: Flowchart

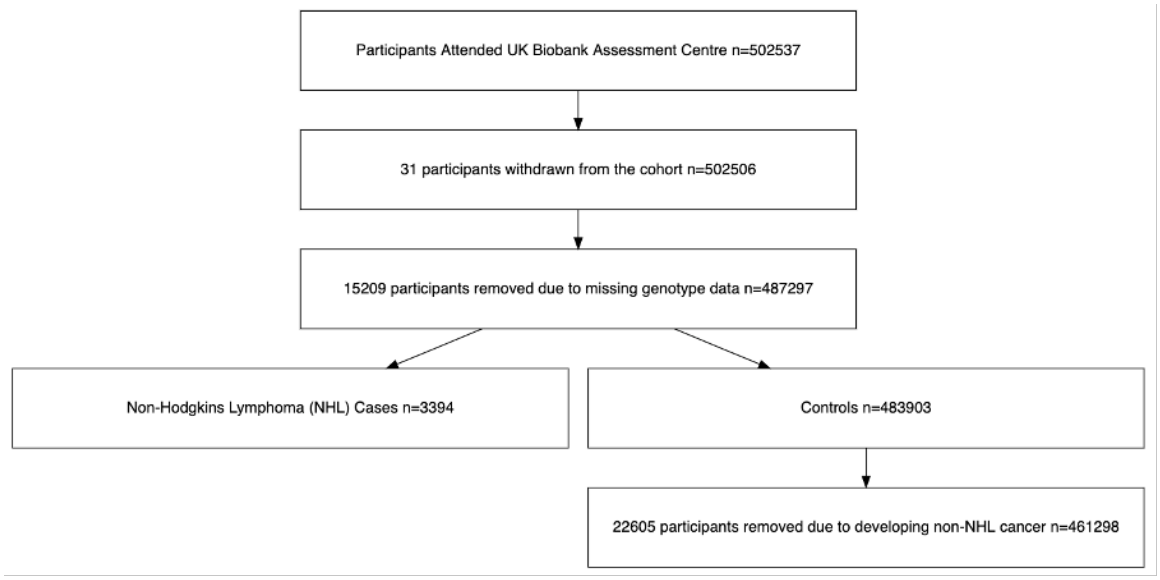


Figure 7: Pcs

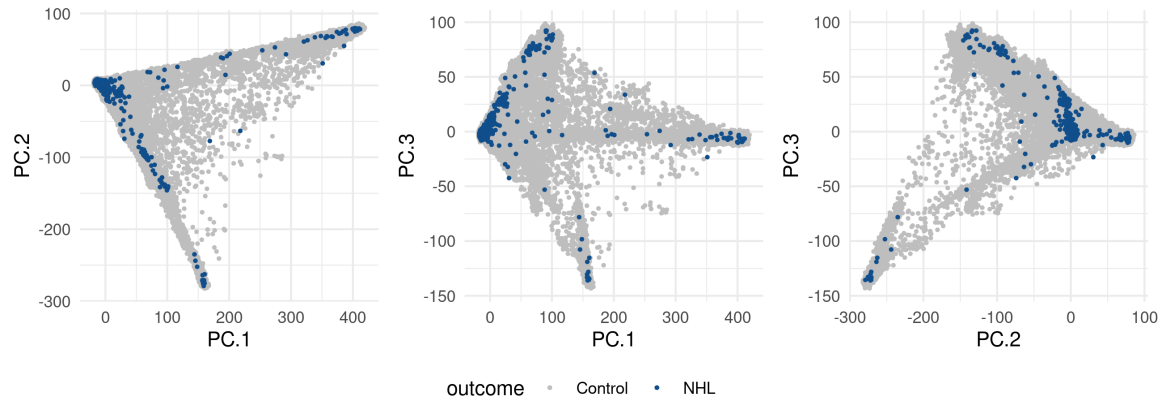
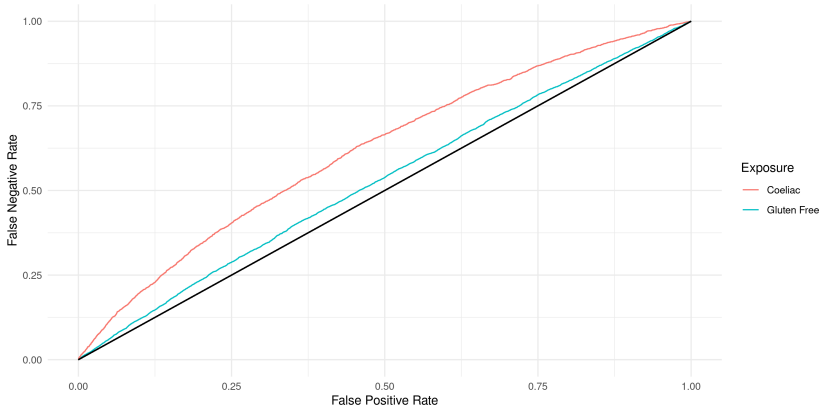


Figure 8: PRS ROC + Table



(a) PRS ROC for Coeliac diagnosis

A	B	C
1	2	3
4	5	6

(b) Blahhh

Table 3: Descriptive 3

	Controls	NHL Cases
n	461298	3394
Male (%)	204230 (44.3)	1861 (54.8)
Age (year)		
Min	49	50
Median (IQR)	69 (13)	74 (9)
Max	86	83
Coeliac (%)	2836 (0.6)	37 (1.1)
Immunodeficient (%)	1795 (0.4)	193 (5.7)
HIV (%)	189 (0.0)	9 (0.3)

Table 4: Marquez SNPs

rsid	Effect Allele	Minor Allele	MAF	p-value	Odds Ratio
rs10469840	T	T	0.258	2.28×10^{-16}	1.19
rs10892299	T	T	0.171	1.97×10^{-11}	0.85
rs10912267	A	A	0.167	1.50×10^{-11}	0.84
rs11712165	G	G	0.417	4.42×10^{-10}	1.12
rs1250568	C	C	0.381	1.13×10^{-14}	0.87
rs12619531	G	G	0.453	4.15×10^{-19}	1.18
rs13132308	G	G	0.126	3.03×10^{-36}	0.71
rs1359062	C	C	0.142	3.50×10^{-30}	0.75
rs17753641	G	G	0.144	6.76×10^{-29}	1.38
rs1893592	C	C	0.255	8.75×10^{-11}	0.88
rs2030519	G	G	0.422	5.38×10^{-52}	0.76
rs212407	G	G	0.393	1.26×10^{-11}	1.13
rs2542148	C	C	0.180	8.34×10^{-12}	1.18
rs3190930	T	T	0.276	1.79×10^{-20}	1.21
rs41432345	C	C	0.104	2.45×10^{-19}	1.35
rs4560096	G	G	0.423	1.56×10^{-13}	1.15
rs60600003	G	G	0.117	3.86×10^{-13}	1.24
rs61907765	T	T	0.242	2.04×10^{-13}	1.17
rs66534072	G	G	0.213	1.07×10^{-9}	1.16
rs6664969	A	A	0.306	1.34×10^{-8}	0.90
rs6691768	G	G	0.354	1.02×10^{-9}	0.89
rs7104791	T	T	0.230	7.75×10^{-10}	1.14
rs7426056	A	A	0.267	1.52×10^{-11}	1.15

Table 5: UKB fields

Field	Field ID	Sub Field	Sub Code	Notes
Sex	31			
Year of Birth	34			Age derived as 2020 - Year of Birth
Type of special diet followed	20086			
		Gluten-free	8	
		Lactose-free	9	
		Low Calorie	10	
		Vegetarian	11	
		Vegan	12	
		Other	13	