BMJ Open Gastroenterology

Investigating the role of iron status in the development of coeliac disease: a Mendelian randomisation study

Isabel A Hujoel [1], Margaux Louise Anna Hujoel [2,3]

To cite: Hujoel IA, Hujoel MLA. Investigating the role of iron status in the development of coeliac disease: a Mendelian randomisation study. *BMJ Open Gastroenterol* 2024;**11**:e001236. doi:10.1136/bmjgast-2023-001236

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjgast-2023-001236).

Received 26 August 2023
Accepted 11 December 2023

[1]University of Washington, Seattle, Washington, USA
[2]Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA
[3]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

Correspondence to
Dr Isabel A Hujoel;
isabelh@uw.edu

ABSTRACT

Objective The environmental trigger behind the increasing prevalence of coeliac disease is not known. One suggested cause is iron deficiency, which is common in coeliac disease. We aimed to evaluate this possible association with Mendelian randomisation (MR), which under certain assumptions can suggest a causal relationship.

Design We conducted a two-sample MR study examining the relationship between single nucleotide polymorphisms (SNPs) associated with iron status and the presence of coeliac disease. The SNPs were drawn from a meta-analysis of three genome-wide association studies (GWAS). The association between these SNPs and coeliac disease was assessed using GWAS summary statistics from the UK Biobank. This consists of 336 638 white British individuals, 1855 with coeliac disease. We performed an MR Egger test for pleiotropy and assessed the plausibility of the assumptions of MR to evaluate for possible causality.

Results There were four SNPs strongly associated with systemic iron status. These were not associated with known risk factors for coeliac disease. All four SNPs were available in the UK Biobank coeliac disease summary statistics. Harmonising exposure and outcome associations, we found that higher iron status was negatively associated with risk of coeliac disease (OR per 1 SD increase in serum iron: 0.65, 95% CI 0.47 to 0.91). Leave-one-out analyses had consistent results, and no single SNP drove the association. All three assumptions of MR appeared plausible.

Conclusion We found that genetically lower iron levels were associated with an increased risk of coeliac disease. Our findings highlight a potential opportunity for coeliac disease prevention.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ There are several postulated environmental triggers for coeliac disease, however evidence behind these is inconclusive, and at this time the reason coeliac disease is increasing in prevalence is not known.

WHAT THIS STUDY ADDS

⇒ Using a study method that can suggest causality, we found that genetically low iron levels are associated with increased coeliac disease risk.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ These findings suggest that iron supplementation in select individuals may provide a potential protective effect against coeliac disease development.

INTRODUCTION

Coeliac disease is an immune-mediated condition, whose development depends on the presence of genetic factors, most notably the haplotypes DQ2 and DQ8, and environmental triggers, crucially gluten intake. Coeliac disease has a world-wide prevalence ranging from 0.7% to 1.8%.[1] The incidence of coeliac disease has been increasing over the past several decades, with one systematic review and meta-analysis estimating a pooled average of a 7.5% increase per year over this time frame.[2] While increased recognition and detection of the disease may be contributing to this rise, there also appears to be a true increase in disease prevalence. The reasons behind this increase are not known, but are thought to be secondary to environmental exposure.[3] Identifying these exposures is paramount, as it offers the opportunity to potentially mitigate this rise.

There have been many environmental triggers proposed for coeliac disease, including breastfeeding duration, timing and quantity of gluten exposure, perinatal factors, infections (including rotavirus), socioeconomic factors, geographical location, microbiome composition and medications.[4–14] The evidence behind these triggers often is based on observational studies, has conflicting results and is not borne out in subsequent interventional studies aimed at coeliac disease prevention.

Iron status has been proposed as an environmental trigger for coeliac disease due to an association between maternal iron supplementation during pregnancy and subsequent coeliac disease in the child, as well as increased coeliac disease in those with hemochromatosis.[15–17] This has been suggested to

be secondary to the impact of iron on the innate immune system and the microbiome.[18 19] We aimed to evaluate the relationship between iron status and coeliac disease through the use of Mendelian randomisation (MR). MR is a statistical method that capitalises on the random allocation of single nucleotide polymorphisms (SNPs) at conception. Through this random allocation and the use of genetic variants which are inherently non-modifiable as instruments for modifiable exposures, MR can avoid confounding factors and can suggest causality as long as certain assumptions hold. These assumptions specify that the instrument must be associated with the exposure, that the instrument does not impact the outcome outside of the exposure of interest, and that the instrument does not influence other potential exposures which may have an impact on the outcome of interest. Crucially for this study, by relying on SNPs associated with iron status, and not on iron status itself, MR may avoid the confounding impact of coeliac disease causing iron deficiency.

METHODS

We conducted a two-sample MR study in which the associations between the genetic instruments (SNPs) and iron status and between the genetic instruments and coeliac disease were measured in different studies. There are three assumptions for the results from this MR study to be valid: (1) the genetic instruments must be associated with systemic iron status (relevance assumption), (2) the genetic instruments must influence coeliac disease only through their effect on systemic iron status (exclusion restriction) and (3) the genetic instruments must not be associated with any confounders, measured or unmeasured (independence assumption).[20]

The exposure of interest was systemic iron status, for which no single biomarker exists but which can be assessed using serum iron, ferritin, transferrin and transferrin saturation biomarkers. As previous studies have done, selected genetic instruments were associated with all four of these biomarkers in a manner consistent with an effect on systemic iron status.[21 22] A recent meta-analysis of three genome-wide association studies (GWASs; from Iceland, the UK and Denmark) identified four such SNPs: *rs1800562* and *rs1799945* in the hemochromatosis (*HFE*) gene, *rs855791* in the transmembrane protease serine 6 (*TMPRSS6*) gene and *rs57659670* predicted to affect the Dual Oxidase 2 (*DUOX2*) gene; the three former instruments were found in a previous iron GWAS whereas the latter variant is novel to this recent meta-analysis.[23 24] We used publicly available GWAS summary statistics from the aforementioned meta-analysis (see the Data availability section). All biomarkers were rank-based inverse normal transformed to a standard normal distribution (on a sex-specific basis) and adjusted for age (within the UK cohort, biomarkers were also adjusted for menopausal status, ABO blood group, body mass index, smoking levels, alcohol levels and iron supplementation status). All associations in the meta-analysis had consistent effects directions across all cohorts. Further details on the meta-analysis methodology and heterogeneity between cohort-specific results are available in Bell *et al*.[24]

The outcome of interest was coeliac disease; we used publicly available GWAS summary statistics from the UK Biobank to assess the association between the genetic instruments for systemic iron status and coeliac disease.[25] These summary statistics result from an analysis performed on data from white individuals of British ancestry; the model was adjusted for genetic relatedness, sex, birth year, and the first four principal components (see the Data availability section). The UK Biobank enrolled roughly 500 000 individuals (aged 40–69 when recruited) between 2006 and 2010. The summary statistics used were restricted to 336 638 white British individuals, of whom 1855 had coeliac disease.[26] The diagnosis of coeliac disease was based on PheCode 557.1 and the summary statistics were computed with SAIGE, which uses a saddlepoint approximation to control for this case–control imbalance, thereby providing accurate p values even when the prevalence of a trait is sufficiently low.[27] PheCodes are a phenotyping tool based on International Classification of Diseases codes.

Both the exposure and outcome studies were conducted in Northern Europe (Iceland, the UK, and Denmark and the UK, respectively). UK samples in the exposure study are from the INTERVAL trial whereas those from the outcome study are from UK Biobank; as both studies are from the UK, sample overlap could exist, however, if it does exist it is likely small and moreover, the INTERVAL study is only one of three studies included in the meta-analysis underlying the exposure data.

To assess the plausibility of the assumption that the genetic instruments only influence coeliac disease through their effect on systemic iron status (exclusion restriction), we searched an online database of SNP-phenotype associations (PhenoScanner, http://www.phenoscanner.medschl.cam.ac.uk/phenoscanner) for secondary phenotypes associated with the four selected instruments and conducted an MR Egger test for pleiotropy.[20 28]

To assess the plausibility of the assumption the genetic instruments are not associated with any confounders (independence assumption) we examined the association between our genetic instruments for iron status and known risk factors for coeliac development available in the UK Biobank, namely sex and age.[29]

As noted previously, no single biomarker exists for systemic iron status and as such we used serum iron levels to quantify the genetic associations of the instruments with systemic iron status as was previously done.[21] To combine information from multiple independent genetic variants into a single causal estimate, we primarily used an inverse-variance weighted (IVW) method; we used random effects when we were analysing at least four variants. An IVW estimate is a consistent estimator if all genetic variants satisfy the three MR assumptions.[30]

Table 1 Harmonised exposure and outcome relationships

		Exposure (serum iron levels)				Outcome (coeliac disease)			
SNP	Effect:other	EAF	Beta	SE	P value	EAF	Beta	SE	P value
rs1799945	G:C	0.1370	0.170	0.00582	1.26×10^{-187}	0.1510	−0.1550	0.0464	0.000842
rs1800562	A:G	0.0677	0.270	0.00760	3.66×10^{-276}	0.0778	−0.0393	0.0618	0.525000
rs57659670	C:T	0.0753	−0.042	0.00735	1.08×10^{-08}	0.0750	0.0360	0.0628	0.566000
rs855791	A:G	0.4310	−0.170	0.00459	1.00×10^{-300}	0.4399	0.0627	0.0335	0.061700

beta, effect size (units of SD for exposure; log of the OR for outcome); EAF, effect allele frequency; effect:other, effect allele:other allele; p value, p value for the association; SNP, single nucleotide polymorphism.

We conducted four sensitivity analyses. First, we ran leave-one-out tests in which the overall association was recomputed when each SNP was left out in turn to ensure no single variant drove the association. Second, we obtained an MR estimate through a weighted median estimator (WME), which is a consistent estimator if at least 50% of the weight comes from variants that satisfy the assumptions.[30] Third, we obtained an MR estimate through MR-Egger which, under a weaker assumption, provides a valid test of the null causal hypothesis and a consistent estimate even if all variants failed to satisfy the three MR assumptions.[31] Fourth, we investigated using the other three biomarkers (other than serum iron levels) to quantify the genetic associations of the instruments with systemic iron status.

All statistical analyses were conducted using the R package MendelianRandomization and MR Base.

Patient and public involvement

Patient and public involvement was not present in this study.

RESULTS

There were four independent and strongly associated SNPs for systemic iron status: *rs1800562* and *rs1799945* in the *HFE* gene, *rs855791* in the *TMPRSS6* gene and *rs57659670* predicted to affect the *DUOX2* gene.[24] Associations between these instruments with the four relevant biomarkers are shown in online supplemental table 1. The two variants in *HFE* are in low linkage disequilibrium (LD) (r^2<0.01) when considering European populations in 1000 genomes using the LDlink resource.[32] We assessed the strength of the instruments through the F-statistic; all four instruments have high F-statistics (F>10; online supplemental table 2).

The associations for all the genetic instruments for systemic iron status were available in the UK Biobank summary statistics; we extracted the summary statistics for these SNPs and harmonised exposure and outcome associations (table 1; online supplemental table 3).

Using PhenoScanner, we found that the genetic instruments were not significantly associated with known risk factors for coeliac disease (online supplemental table 4).[28] There was no significant evidence of directional pleiotropy as the MR-Egger intercept did not differ from the null (p=0.46; table 2). None of the SNPs were associated with either sex or age (online supplemental table 5).

We found that higher iron status was negatively associated with risk of coeliac disease (OR per 1 SD increase in serum iron: 0.65, 95% CI 0.47 to 0.91, p=0.01; table 2, figure 1A). SNP-specific effect associations (ORs) of serum iron with coeliac disease are shown in figure 1A. There was little heterogeneity as assessed with the Cochran Q statistic (Q=4.8, p=0.18).

The results from leave-one-out analyses generally were consistent and there did not appear to be a single variant driving the association (figure 1B). We further note that when both *HFE* related SNPs are excluded results appear consistent (see online supplemental note). The MR estimate obtained through the WME approach was consistent with the primary IVW estimate (OR=0.71, 95% CI 0.52 to 0.98, p=0.035; table 2), while the estimate obtained with MR-Egger was non-significant (OR=0.95, 95% CI 0.34 to 2.7, p=0.92; table 2). However, it has been previously noted that MR-Egger is substantially less efficient than IVW or a WME.[30] We primarily used serum iron levels

Table 2 Mendelian randomisation estimates of the association of systemic iron status (quantified by serum iron levels) with coeliac disease

						MR-Egger	
MR method	OR	95% CI	P value	Cochran Q statistic	Cochran Q statistic p value	Intercept p value	I_{gx}^2, %
IVW	0.65	0.47 to 0.91	0.010	4.8	0.18		
Weighted median	0.71	0.52 to 0.98	0.035				
MR-Egger	0.95	0.34 to 2.7	0.920			0.46	100

IVW, inverse-variance weighted; MR, Mendelian randomisation.

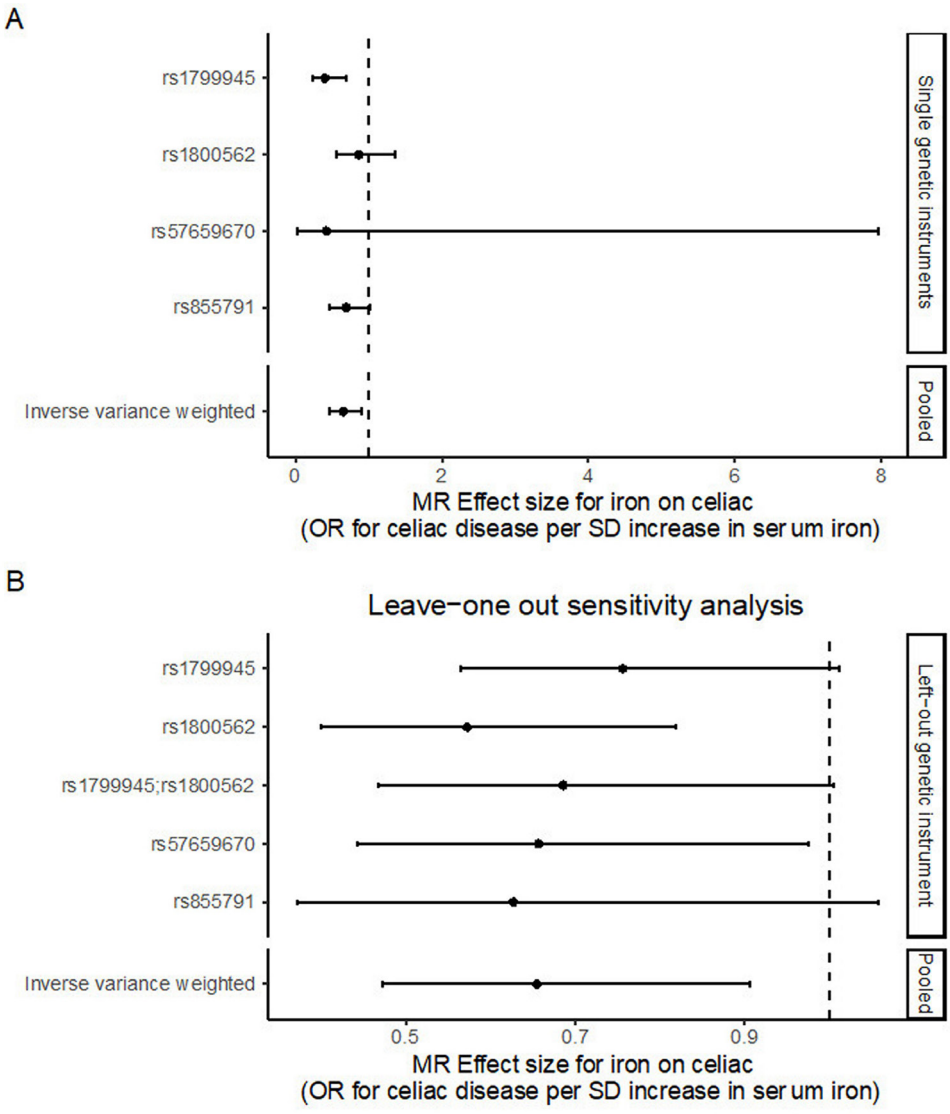

Figure 1 Forest plot of (A) the genetic instrument-specific and summary estimate for the causal effect of systemic iron status (quantified by serum iron levels) on coeliac disease risk, and (B) the summary estimate for the causal effect of systemic iron status on coeliac disease risk leaving each instrument out one at a time. Dots denote point estimates, horizontal lines denote 95% CIs. MR, Mendelian randomisation.

to quantify the genetic associations of the instruments with systemic iron status; analyses using the other three biomarkers are shown in online supplemental table 6 however these analyses were limited by low power (see Supplementary Note).

DISCUSSION

We found that genetically higher iron levels were associated with a decreased risk of having coeliac disease and that, conversely, genetically lower iron levels were associated with a higher risk of having coeliac disease. If the assumptions of MR hold, this association suggests a causal effect of iron deficiency on subsequent coeliac disease development. While iron deficiency is very unlikely to be the only environmental trigger for coeliac disease, iron deficiency is highly prevalent in coeliac disease, and there are intriguing temporal and demographic associations between the two conditions and potential pathophysiological mechanisms.

Iron deficiency is the most common nutritional deficiency globally, and is common in both less-developed and developed countries. In the USA, in fact, iron deficiency has been increasing over the same time period that coeliac disease has.[2 33] This increase in iron deficiency has been attributed to decreased iron intake secondary to both a decrease in iron concentration in food items, and a shift away from iron-rich beef and towards poultry. Iron deficiency is highest in those sex and age-groups where coeliac disease has highest prevalence, such as children and adolescent and premenopausal women.[34 35]

Iron deficiency is highly prevalent in coeliac disease, described in up to roughly 80% of cases, and there has been increasing data to suggest that the relationship between these two conditions is complex.[36] Classically,

iron deficiency in coeliac disease has been attributed to poor iron absorption secondary to the mucosal atrophy in the duodenum, which is the primary site of iron absorption. However, only 10% of dietary iron needs to be absorbed by the duodenum in order to fulfil the body's daily iron needs. Therefore, unless someone's diet is very poor in iron, reduced mucosal surface is likely not to be the only explanation for deficiency. Several studies have also found increased iron deficiency in potential coeliac disease—individuals with elevated coeliac serology but no mucosal damage.[37] After 2 years on a gluten-free diet, another study found that nearly one-half of those with iron deficiency at diagnosis remained iron deficient. In children, those with anaemia had more severe duodenal damage and higher serologic markers, however they were not more likely to have diarrhoea and failure to thrive, and despite a year of a gluten-free diet, they continued to have significantly lower haemoglobin levels than non-anaemic controls.[38 39] Although gastrointestinal bleeding has been proposed as an explanation for these inconsistencies between duodenal damage and iron deficiency, data in support of this theory is controversial.[40 41] This has prompted several studies into potential alterations of iron metabolism regulatory proteins in coeliac disease.[42–46]

One regulatory protein, the transferrin receptor 1, may provide a direct pathogenic role of iron deficiency in coeliac disease development. The transferrin receptor is used to import iron into cells, and in the setting of iron deficiency, it is upregulated, and expression on enterocytes is increased. Studies have demonstrated this increased expression in both treated and untreated coeliac disease patients with iron deficiency.[43 47] Matysiak-Budnik *et al* demonstrated that this receptor also acts to transport un-degraded gliadin peptides across enterocytes. Intestinal transport of intact gliadin peptides to the subepithelial region is a key step in the pathophysiology of coeliac disease.[48 49] The authors note that epithelial cell proliferation (as may be seen following epithelial damage from intestinal infection) also promotes increased expression of transferrin receptor 1, providing an explanation for a possible pathogenic role for gastroenteritis as well as iron deficiency. Soluble transferrin receptor levels are markers of the expression of transferrin receptor in tissue, and are increased in the setting of iron deficiency. Interestingly, one study found that the levels, and thus the expression of transferrin receptor in tissue, was significantly higher in those with coeliac disease and iron deficiency as compared with those with iron deficiency and no coeliac disease. This was thought to be due to the proliferation of epithelial cells in untreated coeliac disease.[50] Iron has also been proposed to contribute to coeliac disease development through its impact on the innate immune system, infection risk and microbiome.[15 18 51 52] Hepcidin and lactoferrin, both proteins involved in iron metabolism, are involved in innate immunity, and iron status has been demonstrated to impact the microbiome, which has been implicated as having a role in the pathogenesis of coeliac disease. Recently, iron has been identified as playing a critical part in regulating the immune system in the intestine, and one recent study suggests that because of this, iron deficiency may play a role in the development of inflammatory bowel disease and other immunological disorders.[53] Coeliac disease may be one of these disorders.

The role of iron status in coeliac disease development has been previously implicated by the increased risk of coeliac disease in those with hereditary hemochromatosis, and in children whose mothers took iron supplementation during pregnancy. Hereditary hemochromatosis is a genetic condition that leads to iron overload. The increased risk for coeliac disease extends beyond 5 years after hemochromatosis diagnosis, when theoretically treatment has been started to decrease iron stores.[17] Indeed, several studies have suggested that the association between the two conditions may be genetic in origin, and that HFE gene mutations (which are common in coeliac disease) may provide a survival advantage in coeliac disease by reducing iron deficiency (Supplementary Note).[54] Notably, in hereditary hemochromatosis, the body does not appear to recognise true iron stores, and soluble transferrin receptor levels are elevated up to a transferrin saturation of 50% (whereas in controls they are elevated only to around 30%).[55] This suggests increased transferrin receptor 1 levels even at higher iron levels in hereditary hemochromatosis and may suggest an additional explanation for the increased coeliac disease risk in the condition. The mechanism behind the increased risk for coeliac disease in children whose mothers took iron supplementation in pregnancy is similarly unclear. This increased risk was not seen in mothers with anaemia or who ate iron rich foods, but was seen in children who took iron supplementation before 18 months. Prior studies have found that maternal iron deficiency leads to increased transferrin receptor expression in the placenta but not soluble transferrin receptor levels in the cord blood.[56 57] It is unknown if transferrin receptors in the placenta can transport intact gliadin peptides across the membrane, similar to enterocytes, theoretically delivering these peptides to the fetus.

This study has several strengths. One strength of this MR study is that it is less likely to be affected by confounding compared with observational studies and, under certain assumptions, suggests a causal relationship. Another strength is that rather than using one genetic instrument, the evidence from multiple instruments were combined, thus providing more power. Moreover, by using a recent meta-analysis of iron GWAS the number of instruments for systemic iron status could be increased from 3 to 4.

This study also has potential limitations. Those with coeliac disease were identified by PheCode, which is derived from international classification of disease codes, and therefore this could have led to misclassification. Additionally, while MR provides some protection against biases such as reverse causation, it is not completely invulnerable.[58] However, we note the mechanism linking the genetic instruments to serum iron

levels (and iron status) is well-understood for most instruments, thereby mitigating the risk of reverse causation. In more detail, if the mechanisms linking the variants to iron status were unclear, it would be possible these variants could influence risk of coeliac disease and cause a downstream effect on iron status, however most of our instruments affect genes known to influence iron status (*HFE* and *TMPRSS6*). Another limitation is that by restricting to instruments that associate with all four iron status markers (with association directions consistent with an effect on systemic iron status), we have the potential to lose power by excluding variants associating with any iron marker.

This study suggests that iron deficiency may play a role in coeliac disease development. A causal relationship should ideally be confirmed with clinical trials, and if affirmative, should be followed by prevention studies. More broadly, this study raises questions of the potential ties between iron deficiency and gluten-sensitivity, and the potential implications of dietary restrictions and avoidant restrictive food intake disorders which are common in coeliac disease, inflammatory bowel disease, and irritable bowel syndrome.

Twitter Isabel A Hujoel @IsabelHujoel

Acknowledgements The findings of this paper were presented at the American College of Gastroenterology 2022 Annual Meeting and were published in abstract form 59.

Contributors Guarantor: IAH. IAH: conceptualisation; writing- original draft. MLAH: formal analysis; data curation; methodology; writing – review and editing.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Ethics approval was not obtained as publicly available data was used.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Summary statistics for the association of SNPs with iron status: Supplementary Data 2 https://www.nature.com/articles/s42003-020-01575-z; https://www.decode.com/summarydata/. Summary statistics for the association of SNPs with celiac disease: UK-Biobank Single Variant Association Analysis Results; https://www.leelabsg.org/resources. Code available on request.

ORCID iD

Isabel A Hujoel http://orcid.org/0000-0002-4065-539X

REFERENCES

1 Singh P, Arora A, Strand TA, *et al*. Global prevalence of celiac disease: systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2018;16:823–36.
2 King JA, Jeong J, Underwood FE, *et al*. Incidence of celiac disease is increasing over time: A systematic review and meta-analysis. *Am J Gastroenterol* 2020;115:507–25.
3 Lebwohl B, Rubio-Tapia A. Epidemiology, presentation, and diagnosis of celiac disease. *Gastroenterology* 2021;160:63–75.
4 Caminero A, McCarville JL, Galipeau HJ, *et al*. Duodenal bacterial proteolytic activity determines sensitivity to dietary antigen through protease-activated Receptor-2. *Nat Commun* 2019;10:1198.
5 Norris JM, Barriga K, Hoffenberg EJ, *et al*. Risk of celiac disease Autoimmunity and timing of gluten introduction in the diet of infants at increased risk of disease. *JAMA* 2005;293:2343–51.
6 Mårild K, Ye W, Lebwohl B, *et al*. Antibiotic exposure and the development of Coeliac disease: a nationwide case-control study. *BMC Gastroenterol* 2013;13:109.
7 Canova C, Zabeo V, Pitter G, *et al*. Association of maternal education, early infections, and antibiotic use with celiac disease: a population-based birth cohort study in northeastern Italy. *Am J Epidemiol* 2014;180:76–85.
8 Lebwohl B, Spechler SJ, Wang TC, *et al*. Use of proton pump inhibitors and subsequent risk of celiac disease. *Dig Liver Dis* 2014;46:36–40.
9 Stene LC, Honeyman MC, Hoffenberg EJ, *et al*. Rotavirus infection frequency and risk of celiac disease Autoimmunity in early childhood: a longitudinal study. *Am J Gastroenterol* 2006;101:2333–40.
10 Koletzko S, Lee H-S, Beyerlein A, *et al*. Cesarean section on the risk of celiac disease in the offspring: the Teddy study. *J Pediatr Gastroenterol Nutr* 2018;66:417–24.
11 Mårild K, Stephansson O, Montgomery S, *et al*. Pregnancy outcome and risk of celiac disease in offspring: a nationwide case-control study. *Gastroenterology* 2012;142:39–45.
12 Ludvigsson JF, Lebwohl B. Three papers indicate that amount of gluten play a role for celiac disease - but only a minor role. *Acta Paediatr* 2020;109:8–10.
13 Celdir MG, Jansson-Knodell CL, Hujoel IA, *et al*. Latitude and celiac disease prevalence: A meta-analysis and meta-regression. *Clin Gastroenterol Hepatol* 2022;20:e1231–9.
14 Vriezinga SL, Auricchio R, Bravi E, *et al*. Randomized feeding intervention in infants at high risk for celiac disease. *N Engl J Med* 2014;371:1304–15.
15 Lebwohl B, Ludvigsson JF, Green PHR. Editorial: the unfolding story of celiac disease risk factors. *Clin Gastroenterol Hepatol* 2014;12:632–5.
16 Stødal K, Haugen M, Brantsæter AL, *et al*. Association between maternal iron supplementation during pregnancy and risk of celiac disease in children. *Clin Gastroenterol Hepatol* 2014;12:624–31.
17 Ludvigsson JF, Murray JA, Adams PC, *et al*. Does Hemochromatosis Predispose to celiac disease? A study of 29,096 celiac disease patients. *Scand J Gastroenterol* 2013;48:176–82.
18 Nairz M, Schroll A, Sonnweber T, *et al*. The struggle for iron - a metal at the host-pathogen interface. *Cell Microbiol* 2010;12:1691–702.
19 Zimmermann MB, Chassard C, Rohner F, *et al*. The effects of iron Fortification on the gut Microbiota in African children: a randomized controlled trial in Cote D'Ivoire. *Am J Clin Nutr* 2010;92:1406–15.
20 Davies NM, Holmes MV, Davey Smith G. Reading Mendelian Randomisation Sudies: a guide, glossary, and checklist for Clinicians. *BMJ* 2018;362:k601.
21 Gill D, Benyamin B, Moore LSP, *et al*. Associations of genetically determined iron status across the Phenome: A Mendelian randomization study. *PLoS Med* 2019;16:e1002833.
22 Gill D, Del Greco M F, Walker AP, *et al*. The effect of iron status on risk of coronary artery disease: A Mendelian randomization study-brief report. *Arterioscler Thromb Vasc Biol* 2017;37:1788–92.
23 Benyamin B, Esko T, Ried JS, *et al*. Novel Loci affecting iron homeostasis and their effects in individuals at risk for Hemochromatosis. *Nat Commun* 2014;5:4926.
24 Bell S, Rigas AS, Magnusson MK, *et al*. A genome-side meta-analysis yields 46 new Loci associating with biomarkers of iron homeostasis. *Commun Biol* 2021;4:156.
25 Bycroft C, Freeman C, Petkova D, *et al*. The UK Biobank resource with deep Phenotyping and Genomic data. *Nature* 2018;562:203–9.
26 Lee lab. UK-Biobank single variant association analysis results. 2018. Available: https://www.leelabsg.org/resources
27 Zhou W, Nielsen JB, Fritsche LG, *et al*. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic Association studies. *Nat Genet* 2018;50:1335–41.

28 Staley JR, Blackshaw J, Kamat MA, *et al*. Phenoscanner: a database of human genotype-phenotype associations. *Bioinformatics* 2016;32:3207–9.
29 GWAS Results. Neale lab. n.d. Available: http://www.nealelab.is/uk-biobank
30 Bowden J, Davey Smith G, Haycock PC, *et al*. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median Estimator. *Genet Epidemiol* 2016;40:304–14.
31 Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512–25.
32 Machiela MJ, Chanock SJ. Ldlink: a web-based application for exploring population-specific haplotype structure and linking correlated Alleles of possible functional variants. *Bioinformatics* 2015;31:3555–7.
33 Sun H, Weaver CM. Decreased iron intake parallels rising iron deficiency anemia and rleated mortality rates in the US population. *J Nutr* 2021;151:1947–55.
34 Martín-Masot R, Nestares MT, Diaz-Castro J, *et al*. Multifactorial etiology of anemia in celiac disease and effect of gluten-free diet: A comprehensive review. *Nutrients* 2019;11:2557.
35 Turawa E, Awotiwon O, Dhansay MA, *et al*. n.d. Prevalence of anaemia, iron deficiency, and iron deficiency anaemia in women of reproductive age and children under 5 years of age in South Africa (1997-2021): A systematic review. *IJERPH*;18:12799.
36 Berry N, Basha J, Varma N, *et al*. Anemia in celiac disease is Multifactorial in etiology: A prospective study from India. *JGH Open* 2018;2:196–200.
37 Repo M, Lindfors K, Mäki M, *et al*. Anemia and iron deficiency in children with potential celiac disease. *J Pediatr Gastroenterol Nutr* 2017;64:56–62.
38 Rajalahti T, Repo M, Kivelä L, *et al*. Anemia in pediatric celiac disease: association with clinical and histological features and response to gluten-free diet. *J Pediatr Gastroenterol Nutr* 2017;64:e1–6.
39 Grisolano SW, Oxentenko AS, Murray JA, *et al*. The usefulness of routine small bowel biopsies in evaluation of iron deficiency anemia. *J Clin Gastroenterol* 2004;38:756–60.
40 Mant MJ, Bain VG, Maguire CG, *et al*. Prevalence of occult gastrointestinal bleeding in celiac disease. *Clin Gastroenterol Hepatol* 2006;4:451–4.
41 Logan RFA, Howarth GF, West J, *et al*. How often is a positive Faecal occult blood test the result of Coeliac disease *Eur J Gastroenterol Hepatol* 2003;15:1097–100.
42 Repo M, Hannula M, Taavela J, *et al*. Iron transporter protein expressions in children with celiac disease. *Nutrients* 2021;13:776.
43 Barisani D, Parafioriti A, Bardella MT, *et al*. Adaptive changes of Duodenal iron transport proteins in celiac disease. *Physiol Genomics* 2004;17:316–25.
44 Sharma N, Begum J, Eksteen B, *et al*. Differential Ferritin expression is associated with iron deficiency in coeliac disease. *Eur J Gastroenterol Hepatol* 2009;21:794–804.
45 De Falco L, Tortora R, Imperatore N, *et al*. The role of Tmprss6 and HFE variants in iron deficiency anemia in celiac disease. *Am J Hematol* 2018;93:383–93.
46 Tolone C, Bellini G, Punzo F, *et al*. The Dmt1Ivs4+44C>A polymorphism and the risk of iron deficiency anemia in children with celiac disease. *PLoS One* 2017;12:e0185822.
47 McDevitt J, O'Farrelly C, Weir DG, *et al*. Proliferation-associated markers in coeliac duodenum. *Europ J Gastroenterol Hepatol* 1994;6:223–8.
48 Kagnoff MF. Overview and pathogenesis of celiac disease. *Gastroenterology* 2005;128:S10–8.
49 Matysiak-Budnik T, Moura IC, Arcos-Fajardo M, *et al*. Secretory IgA mediates Retrotranscytosis of intact Gliadin peptides via the Transferrin receptor in celiac disease. *J Exp Med* 2008;205:143–54.
50 De Caterina M, Grimaldi E, Di Pascale G, *et al*. The soluble Transferrin receptor (sTfR)-Ferritin index is a potential Predictor of celiac disease in children with refractory iron deficiency anemia. *Clin Chem Lab Med* 2005;43:38–42.
51 Johnson EE, Wessling-Resnick M. Iron metabolism and the innate immune response to infection. *Microbes Infect* 2012;14:207–16.
52 Rusu IG, Suharoschi R, Vodnar DC, *et al*. Iron supplementation influence on the gut Microbiota and Probiotic intake effect in iron deficiency - A literature-based review. *Nutrients* 2020;12:1993.
53 Xiong L, Helm EY, Dean JW, *et al*. Nutrition impact on Ilc3 maintenance and function centers on a cell-intrinsic Cd71-iron axis. *Nat Immunol* 2023;24:1671–84.
54 Butterworth JR, Cooper BT, Rosenberg WMC, *et al*. The role of Hemochromatosis susceptibility gene mutations in protecting against iron deficiency in celiac disease. *Gastroenterology* 2002;123:444–9.
55 Brandão M, Oliveira JC, Bravo F, *et al*. The soluble Transferrin receptor as a marker of iron homeostasis in normal subjects and in HFE-related Hemochromatosis. *Haematologica* 2005;90:31–7.
56 Sweet DG, Savage G, Tubman TR, *et al*. Study of maternal influences on fetal iron status at term using cord blood Transferrin receptors. *Arch Dis Child Fetal Neonatal Ed* 2001;84:F40–3.
57 Sangkhae V, Fisher AL, Wong S, *et al*. Effects of maternal iron status on Placental and fetal iron homeostasis. *J Clin Invest* 2020;130:625–40.
58 Burgess S, Swanson SA, Labrecque JA. Are Mendelian randomization investigations immune from bias due to reverse causation *Eur J Epidemiol* 2021;36:253–7.

Supplementary Note

***HFE* and celiac disease**

Butterworth et al. (commentary by Bowlus and Lie *Gastroenterology* 2003; Ravine and Darke *Gastroenterology* 2003) suggested that *HFE* gene mutations are common in individuals with celiac disease and are in linkage disequilibrium with different HLA alleles as compared to individuals without celiac disease; it was suggested that *HFE* gene mutations may provide a survival advantage in celiac disease by reducing iron deficiency.

If mutations within *HFE* affect the probability to have celiac disease (independent of their effect on systemic iron status), the exclusion restriction assumption of mendelian randomization is violated. Notably, if variants within *HFE* are in strong LD with various HLA alleles associated with celiac disease, as was suggested (see 2 related commentaries for weaknesses in Butterworth et al. and potential alternative explanations for the observations made); this could result in a spurious association of *HFE* mutations and celiac disease. Specifically, Butterworth et al. studied the *HFE* mutations, C282Y (also known as rs1800562) and H63D (also known as rs1799945), both of which are instruments in our study.

To test whether the genetic instruments within *HFE* drove our association, we ran analyses excluding both instruments within *HFE;* results appeared consistent (Figure 1b), suggesting the SNPs within *HFE* did not drive our association.

Using each of the 4 biomarkers to quantify systemic iron status

We primarily used serum iron levels to quantify the genetic associations of the instruments with systemic iron status; analyses using the other 3 biomarkers are shown in Supplementary Table 5. All I_{GX}^2 statistics are sufficiently large.

When using ferritin levels to quantify the associations, we found significant heterogeneity (Q=8.6, p = 0.036), thereby invalidating the IVW estimate; using the WM estimator we found no significant relationship between iron status and risk of celiac disease (OR: 0.74, 95% CI 0.38-1.5, P = 0.4; Supplementary Table 5). We note that the power to detect an OR of 0.8 using ferritin as the quantifier of associations was low (11% power; Supplementary Table 6).

When using transferrin saturation levels to quantify the associations, which had a higher power to detect an OR of 0.8 (49% power; Supplementary Table 6), we found a significant relationship between iron status and risk of celiac disease (OR: 0.74, 95% CI 0.56-0.99, P = 0.04; Supplementary Table 5).

When using total iron-binding capacity to quantify the associations, we found significant heterogeneity (Q=12, p = 0.009) as well as a significant MR-Egger intercept (P=0.041), thereby invalidating the IVW estimate. Using the WM estimator, we found no significant relationship between iron status and risk of celiac disease (OR: 1.2, 95% CI 0.89-1.5, P = 0.28; Supplementary Table 5). The power to detect an OR of 1.2 using variants that explain 3.01% in variance of the exposure (similar to total iron-binding capacity) is 32% (Supplementary Table 6).

References:

Butterworth J, Cooper B, Rosenberg W, et al. The role of hemochromatosis susceptibility gene mutations in protecting against iron deficiency in celiac disease. *Gastroenterology*. 2002;123:444-449.

Bowlus, C.L. and Lie, B.A.. Discussion on the Role of Hemochromatosis Susceptibility Gene Mutation in Protecting Against Iron Deficiency in Celiac Disease. *Gastroenterology* 2003; 124: 1562 – 1563.

Ravine, D. and Darke, C.. Discussion on the Role of Hemochromatosis Susceptibility Gene Mutation in Protecting Against Iron Deficiency in Celiac Disease. *Gastroenterology* 2003; 124: 1563.

Supplementary Table 1: **Results from Bell et al. 2021**. For all selected genetic instruments, the position, minor and major allele, minor allele frequency, and effect on 4 relevant iron biomarkers is shown. The effect is shown for the minor allele.

Rsid	Position (hg38)	Minor: Major Allele	MAF (%)	Phenotype	Effect in SD (95% CI)	P-value
rs1799945	chr6:26090951	G:C	13.7	TSAT	0.21 (0.2; 0.23)	6.1e-229
				TIBC	-0.12 (-0.13; -0.1)	4.29e-66
				Ferritin	0.059 (0.049; 0.069)	1.51e-31
				Iron	0.17 (0.16; 0.18)	1.26e-187
rs1800562	chr6:26092913	A:G	6.77	TIBC	-0.45 (-0.47; -0.43)	1e-300
				TSAT	0.45 (0.42; 0.47)	1e-300
				Ferritin	0.13 (0.12; 0.15)	1.85e-84
				Iron	0.27 (0.26; 0.29)	3.66e-276
rs57659670	chr15:45106240	C:T	7.53	Ferritin	-0.14 (-0.16; -0.13)	1.05e-113
				Iron	-0.042 (-0.056; -0.028)	1.08e-08
				TIBC	0.077 (0.06; 0.094)	3.67e-19
				TSAT	-0.058 (-0.074; -0.041)	5.73e-12
rs855791	chr22:37066896	A:G	43.1	TSAT	-0.17 (-0.18; -0.16)	1e-300
				Iron	-0.17 (-0.18; -0.16)	1e-300
				TIBC	0.026 (0.017; 0.035)	2.88e-08
				Ferritin	-0.044 (-0.051; -0.038)	6.14e-37

MAF: minor allele frequency; SD: standard deviation; CI: confidence interval; TIBC: total iron-binding capacity; TSAT transferrin saturation.

Supplementary Table 2: **Strength of instruments.** For 4 iron biomarkers, the amount of variance explained by each instrument as well as the F-statistic is computed. As effect sizes are in units of standard deviation, the percentage of variation in iron biomarker explained by the SNP (R^2); we computed this as $2 \times \beta^2 \times AF \times (1 - AF)$. Total R^2 values are 0.65, 3.04, 3.01, and 5.07 for ferritin, iron, TIBC, and TSAT, respectively. F-statistics (F) were computed as a function of the variance explained and sample size (Palmer et al. *Stat Methods Med Res* 2012).

	Effect on exposure (units of SD)				
SNP	Estimate	SE	P-value	R^2	F
Ferritin					
rs1799945	0.059	0.005	1.51e-31	0.08	202.77
rs1800562	0.13	0.0067	1.85e-84	0.21	526.22
rs57659670	-0.14	0.0062	1.05e-113	0.27	673.67
rs855791	-0.044	0.0035	6.14e-37	0.09	233.95
Iron					
rs1799945	0.17	0.0058	1.26e-187	0.68	1125.08
rs1800562	0.27	0.0076	3.66e-276	0.92	1518.66
rs57659670	-0.042	0.0073	1.08e-08	0.02	40.18
rs855791	-0.17	0.0046	1e-300	1.42	2351.05
TIBC (total iron-binding capacity)					

rs1799945	-0.12	0.007	4.29e-66	0.34	462.72
rs1800562	-0.45	0.012	1e-300	2.56	3552.69
rs57659670	0.077	0.0086	3.67e-19	0.08	111.91
rs855791	0.026	0.0047	2.88e-08	0.03	44.92
TSAT (transferrin saturation)					
rs1799945	0.21	0.0065	6.1e-229	1.04	1385.41
rs1800562	0.45	0.012	1e-300	2.56	3448.83
rs57659670	-0.058	0.0084	5.73e-12	0.05	61.62
rs855791	-0.17	0.0046	1e-300	1.42	1890.36

Supplementary Table 3: **Reported exposure and outcome relationships.** For all 4 genetic instruments, the reported effect allele, allele frequency of effect allele (EAF), and association statistics are reported for serum iron levels and celiac disease. Bolded SNP had to be harmonized across datasets.

	Exposure (serum iron levels)					
SNP	Effect allele	Other allele	EAF	Beta	SE	P-val
rs1799945	G	C	0.1370	0.170	0.00582	1.26e-187
rs1800562	A	G	0.0677	0.270	0.00760	3.66e-276
rs57659670	C	T	0.0753	-0.042	0.00735	1.08e-08
rs855791	**A**	**G**	**0.4310**	**-0.170**	**0.00459**	**1.00e-300**
Outcome (celiac disease)						
rs1799945	G	C	0.1510	-0.1550	0.0464	0.000842
rs1800562	A	G	0.0778	-0.0393	0.0618	0.525000
rs57659670	C	T	0.0750	0.0360	0.0628	0.566000
rs855791	**G**	**A**	**0.5601**	**-0.0627**	**0.0335**	**0.061700**

Beta: effect size (units of standard deviation for exposure; log of the odds ratio for outcome); SE standard error; P-val p-value for the association.

Supplementary Table 4: Phenotypes associated with included SNPs as determined through Phenoscanner.

SNP	Associated phenotypes
rs1800562	mean corpuscular hemoglobin; disorders of mineral metabolism; iron status biomarkers; transferrin saturation; red cell distribution width; mean corpuscular hemoglobin concentration; transferrin levels; hemoglobin concentration; iron levels; reticulocyte fraction of red cells; reticulocyte count; hematocrit; mean corpuscular hemoglobin; transferrin; self-reported hereditary or genetic haematological disorder; transferrin saturation with iron; mean corpuscular volume; high light scatter percentage of red cells; ferritin levels; high light scatter reticulocyte count; alcohol consumption; transferrin glycosylation; carbohydrate deficient transferrin supplementary concept; HbA1c; Glycated hemoglobin levels; Hematological parameters; Erythrocyte

	indices; Hemoglobin a glycosylated; Hemoglobin; Iron; Total iron binding capacity; Ferritin; Diastolic blood pressure; Comparative height size at age 10; Transferrin saturation; Hepcidin levels; Red blood cell count; Polycythaemia vera; LDL cholesterol; Low density lipoprotein; Total cholesterol; Cholesterol total; Cardiovascular disease risk factors; Height; Pulse rate; Impedence of arm left; Ferritin log10; Hematocrit; Iron deficiency; Unsaturated iron binding capacity; Vascular or heart problems diagnosed by doctor: high blood pressure; Soluble transferrin receptor; Forced vital capacity; Self-reported hypertension; Platelet count; Cause of death: liver cell carcinoma; Monocyte count; Forced expiratory volume in 1 second; Impedence of arm right; Vascular or heart problems diagnosed by doctor: none of the above; Forced vital capacity, best measure; Hepcidin supplementary concept; Impedence of whole body; Red blood cell traits; Plateletcrit; Other complications of surgical and medical care; Hamatology traits; Forced expiratory volume in 1-second, best measure; Long-standing illness, disability or infirmity; Medication for cholesterol, blood pressure or diabetes: blood pressure medication; Platelet distribution width
rs1799945	mean corpuscular hemoglobin; mean corpuscular volume; red cell distribution width; mean corpuscular hemoglobin concentration; hemoglobin concentration; hematocrit; reticulocyte fraction of red cells; iron status biomarkers transferrin saturation; reticulocyte count; iron status biomarkers iron levels; vascular or heart problems diagnosed by doctor: high blood pressure; self-reported hypertension; HbA1c; vascular or heart problems diagnosed by doctor: none of the above; diastolic blood pressure; blood pressure; treatment with blood pressure medication; platelet count; mean arterial pressure; systolic blood pressure; red blood cell count; medication for cholesterol, blood pressure or diabetes: blood pressure medication; hypertension; iron status biomarkers ferritin levels; high light scatter percentage of red cells; hemoglobin; iron status biomarkers; iron; serum iron; plateletcrit; immature fraction of reticulocytes; treatment with Bendroflumethiazide; no treatment with medication for cholesterol, blood pressure, diabetes, or take exogenous hormones; high light scatter reticulocyte count; illness of siblings: high blood pressure; illness of mother: high blood pressure; treatment with lisinopril; platelet distribution width; serum urate
rs855791	Mean corpuscular hemoglobin; mean corpuscular volume, red cell distribution width; mean corpuscular hemoglobin concentration; hemoglobin concentration; iron status biomarkers iron levels; iron status biomarkers transferrin saturation; hematocrit; red blood cell traits; clinical laboratory measurements; reticulocyte fraction of red cells; hemoglobin; reticulocyte count; HbA1c; erythrocyte indices; transferrin saturation with iron ; iron; plateletcrit; platelet count; immature fraction of reticulocytes; soluble transferrin receptor; iron status biomarkers; iron regulatory proteins; iron status biomarkers ferritin levels; glycated hemoglobin HbA1c; hemoglobin a glycosylated; red blood cell count; hepcidin transferrin saturation ratio; transferrin saturation; hepcidin levels; transferrin; hematology traits; platelet

	distribution width; high light scatter percentage of red cells; mineral and other dietary supplements: iron; ferritin
rs57659670	None

Supplementary Table 5: **Association of genetic instruments with sex and age.**

	Exposure		Outcome			
SNP	Eff: Other	AF	Variant	AF	Sex p-value	Age p-value
rs1799945	G:C	0.1370	6:26091179:C:G	0.1512	0.695	0.732
rs1800562	A:G	0.0677	6:26093141:G:A	0.0788	0.852	0.248
rs57659670	C:T	0.0753	15:45398438:T:C	0.0759	0.556	0.570
rs855791	A:G	0.4310	22:37462936:A:G	0.5615	0.107	0.727

Variant identifier in the form "chr:pos:ref:alt", where "ref" is aligned to the forward strand of GRCh37 and "alt" is the effect allele.

Supplementary Table 6: **MR results when using each of the 4 biomarkers in turn to quantify the genetic associations of the instruments with systemic iron status.** IVW, weighted median, and MR-Egger results for all 4 biomarkers is shown. Highlighted row shows main results (Table 2).

Biomarker	MR Method	OR	95% CI	P	Cochran Q Statistic	Cochran Q Statistic P value	MR-Egger	
							Intercept P value	I_{gx}^2,%
Ferritin	IVW	0.48	(0.19:1.2)	0.12	8.6	0.036		
	Weighted median	0.74	(0.38:1.5)	0.4				
	MR-Egger	1.6	(0.38:7)	0.51			0.061	96
Iron	IVW	0.65	(0.47:0.91)	0.01	4.8	0.18		
	Weighted median	0.71	(0.52:0.98)	0.035				
	MR-Egger	0.95	(0.34:2.7)	0.92			0.46	100
TIBC	IVW	1.3	(0.79:2.1)	0.32	12	0.0091		
	Weighted median	1.2	(0.89:1.5)	0.28				
	MR-Egger	0.96	(0.62:1.5)	0.85			0.041	100
TSAT	IVW	0.74	(0.56:0.99)	0.044	6.5	0.088		
	Weighted median	0.77	(0.61:0.99)	0.038				
	MR-Egger	1	(0.57:1.8)	0.97			0.23	100

CI confidence interval; TIBC total iron-binding capacity; TSAT transferrin saturation.

Supplementary Table 7: **Power analyses**: We conduct various power analyses (Brion et al. *IJE* 2013; https://shiny.cnsgenomics.com/mRnd/) for a variety of R^2 values. We assumed a sample size of 336638, with a case prevalence of 1855/336638, and type-I error rate of 5%. We tested R^2 values of 0.65, 3.04, 3.01, and 5.07 to mirror ferritin, iron, TIBC, and TSAT, respectively.

R^2	OR	Power
0.65 (ferritin)	0.8	0.11
3.04 (iron)	0.8	0.32
5.07 (TSAT)	0.8	0.49
3.01 (TIBC)	1.2	0.32