Serum Iron Level and Kidney Function: a Mendelian Randomization Study

Fabiola Del Greco M[1*] PhD, Luisa Foco[1] PhD, Irene Pichler[1] PhD, Philipp Eller[2] MBA, Kathrin Eller[2] MD, Beben Benyamin[3,4]PhD, John B Whitfield[4] PhD, Genetics of Iron Status Consortium, CKDGen Consortium, Peter P Pramstaller[1] MD, John R Thompson[5] PhD, Cristian Pattaro[1] PhD, Cosetta Minelli[6*] PhD

1. Center for Biomedicine, European Academy of Bolzano/Bozen (EURAC), affiliated to the University of Lübeck, Bolzano, Italy
2. Medical University of Graz, Graz, Austria
3. Queensland Brain Institute, The University of Queensland, Queensland, Australia
4. QIMR Berghofer Medical Research Institute, Brisbane, Australia
5. Department of Health Sciences, University of Leicester, Leicester, UK
6. Respiratory Epidemiology, Occupational Medicine and Public Health, National Heart and Lung Institute, Imperial College, London, United Kingdom

Corresponding authors:

Fabiola Del Greco M, PhD
Center for Biomedicine – EURAC research Bozen/Bolzano
Via Galvani 31
39100 Bolzano, Italy
E-mail: fabiola.delgreco@eurac.edu
Tel: +39 0471 055 528; Fax: +39 0471 055 599

Cosetta Minelli, MD PhD
NHLI, Imperial College London
Emmanuel Kaye Building
1 Manresa Road
SW3 6LR London, United Kingdom
E-mail: cosetta.minelli1@imperial.ac.uk
Tel: +44 (0) 207 5947758

Abstract word count: ~~247~~ 244 words

Total word count: ~~3,289~~ 3,490 words

Running title: Iron and eGFR

Keywords: Mendelian randomization; iron; ferritin; glomerular filtration rate; kidney function; chronic kidney disease

Abstract

Background. Iron depletion is a known consequence of chronic kidney disease (CKD), but there is contradicting epidemiological evidence on whether iron itself affects kidney function and whether its effect is protective or detrimental in the general population. While epidemiological studies tend to be affected by confounding and reverse causation, Mendelian Randomization (MR) can provide unconfounded estimates of causal effects by using genes as instruments.

Methods. We performed a MR study of the effect of serum iron levels on estimated glomerular filtration rate (eGFR), using genetic variants known to be associated with iron. MR estimates of the effect of iron on eGFR were derived based on the association of each variant with iron and eGFR from two large genome-wide meta-analyses on 48,978 and 74,354 individuals, respectively. We performed a similar MR analysis for ferritin, which measures iron stored in the body, using variants associated with ferritin.

Results. A combined MR estimate across all variants showed a 1.3% increase in eGFR per standard deviation increase in iron (95%CI: 0.4-2.1%; *p-value*: 0.004). The results for ferritin were consistent with those for iron. Secondary MR analyses of the effects of iron and ferritin on CKD did not show significant associations but had very low statistical power.

Conclusions. Our study suggests a protective effect of iron on kidney function in the general population. Further research is required to confirm this causal association, investigate it in study populations at higher risk of CKD, and explore its underlying mechanism of action. ~~Further experimental and functional studies to confirm these findings could have important public health implications, given the high frequency of iron deficiency and its simple correction through iron supplementation.~~

Introduction

It is well known that chronic kidney disease (CKD) can cause iron depletion, and iron supplementation is widely used to prevent and treat anaemia in CKD patients [1, 2]. What is not known is whether iron can influence kidney function and disease in the general population, and it is unclear whether altered iron levels could also be a cause, rather than a consequence, of kidney disease. The deleterious effects of iron accumulation in tissues in iron-overload conditions is well recognized, and experimental models of both acute kidney injury and CKD have suggested that iron could be nephrotoxic even in the absence of systemic overload, possibly through the formation of free radicals [3-6]. However, studies in rats have shown that ~~iron deficiency~~ dietary iron elimination ~~too~~ can cause tissue damage through oxidative stress and mitochondrial damage [7, 8]. Cellular iron homeostasis plays an important role in mitochondrial biogenesis, and iron deprived cells show downregulation of mitochondrial protein levels and oxidative capacity [9].

Assessing the effects of iron on kidney function in humans using observational data is very challenging because of the difficulty in disentangling causal from spurious effects due to confounding and reverse causation. On the other hand, randomized clinical trials investigating the effect of iron supplementation on kidney function in healthy individuals would be difficult to justify. In situations like this, Mendelian randomization (MR) is being increasingly used to provide estimates of causal effects of modifiable risk factors (e.g. biomarkers) on disease traits. MR is an instrumental variable approach based on the use of genes as 'proxies' (instruments) for the biomarker: genetic variants known to be associated with altered levels of the biomarker are investigated for their effect on the clinical endpoint [10, 11]. Being carried from birth, genes cannot be influenced by changes in the clinical endpoint (no reverse causation) or life-style and environmental predictors (no classical confounding). Moreover,

the MR approach estimates the lifetime effect of the risk factor on the clinical endpoint, as opposed to the effect at a certain point in life estimated by classical observational studies.

The rapidly accumulating evidence on genetic associations for biomarkers (and other risk factors) and disease traits explains the increasing interest in the use of MR to distinguish between spurious and causal associations. As any other instrumental variable approach, MR suffers from limited statistical power [10], but the availability of very large genetic datasets currently ensures adequate power to investigate the causal effects on disease of many risk factors. However, the validity of MR heavily depends on the validity of the instruments chosen, which makes the selection of the genetic variants the most important step of a MR study. A key aspect is the "strength" of the instrument (the genetic variant), defined as a function of both the magnitude of the genetic effect on the risk factor and the precision with which it is estimated; the use of weak instruments can introduce bias (weak-instrument bias) [12]. Given the modest genetic effects typical of common variants, adequate instrument strength is usually achieved through large sample sizes for the gene-risk factor association. Another fundamental aspect is the absence of pleiotropy, which means that the instrument influences the disease trait only through the risk factor considered and not through any other pathway. This is the most challenging assumption in MR and its violation can produce biased MR estimates [10, 11].

In this study we have used MR to investigate the lifetime causal effect of serum iron levels on kidney function in the general population. Kidney function was evaluated based on the glomerular filtration rate, estimated from serum creatinine using the four-parameter MDRD equation (eGFR). We used genetic data from the largest genome-wide association (GWA) consortia for iron markers (GIS consortium) [13] and eGFR (CKDGen consortium) [14]. As

instruments we selected all variants found associated with serum iron levels and replicated in the GIS consortium, with the exception of one variant which we excluded due to possible pleiotropic effects on kidney. To investigate whether findings for the effect of circulating iron levels on kidney function could be also observed for iron stored in the body, we performed a similar MR analysis for the effect of serum ferritin levels on eGFR, again using all genetic variants associated with ferritin and replicated in the GIS consortium. Finally, we repeated the MR analyses to estimate the effects of both iron and ferritin on the risk of CKD, defined as eGFR<60 ml/min/1.73 m^2.

Methods

Selection of genetic variants and investigation of pleiotropy

We considered as instruments for the MR analyses all genetic variants reaching genome-wide significance ($p<5x10^{-8}$) in the combined discovery and replication analysis of the GIS consortium on serum iron and ferritin, based on a total sample size of 48,978 individuals from 18 studies, which represents the largest GWA meta-analysis conducted so far on these phenotypes [13].

For the MR analysis of the effect of iron on eGFR, five variants were considered: rs855791 (also known as V736A) in the transmembrane protease 6 (*TMPRSS6*) gene; rs1800562 (C282Y) and rs1799945 (H63D) in the hemochromatosis (*HFE*) gene; rs8177240 in the transferrin (*TF*) gene; rs7385804 in the transferrin receptor 2 (*TFR2*) gene. All variants are independent, that is they are not in linkage disequilibrium (LD), including the two variants in *HFE* (LD $r^2=0.01$) located in genomic regions coding for functionally distinct protein domains [15, 16]. These variants were associated with serum iron levels at p-values ranging

from 10^{-20} to 10^{-139}. Three of them, rs855791, rs1800562, and rs1799945 were already known to affect serum iron levels [17, 18].

For the MR analysis of the effect of ferritin on eGFR, six variants were considered, three of which were the same as for iron, *TMPRSS6* rs855791, *HFE* rs1800562, and *HFE* rs1799945; the other three included a variant located between the WD Repeat Domain 75 (*WDR75*) and the Solute Carrier Family 40 (Iron-Regulated Transporter) Member 1 (*SLC40A1*) genes, rs744653, a variant in the ABO Blood Group (*ABO*) gene, rs651007, and a variant in the Testis Expressed 14 (*TEX14*) gene, rs411988. These six variants were associated with ferritin with p-values ranging from 10^{-8} to 10^{-38}.

For all variants for iron and ferritin, we investigated the possible presence of pleiotropy both *a priori* based on biological knowledge, and *a posteriori* based on statistical evidence. For the biological assessment, we performed a literature search to identify evidence on the effects of the genes in which the selected variants are annotated. In cases where the variants were acting as *cis* and/or *trans* expression quantitative trait loci (eQTLs), we also included in our investigation any additional gene showing differential expression. For the latter task, we used several freely available online tools, including GTEx [19], Blood eQTL database [20], and ENCODE project data at UCSC Genome browser [21]. For the statistical assessment of possible unknown pleiotropic effects, or other factors invalidating the MR assumption that any effect of the variant on eGFR is mediated by iron (or ferritin), we investigated the heterogeneity of MR estimates across instruments, using the heterogeneity Q test (*p-het*) and the I^2 index as we have previously described [22].

Estimates for the gene-iron and gene-ferritin associations

The effects of the genetic variants on serum iron and ferritin were based on the GIS GWA meta-analysis on 48,978 individuals from the 18 studies. ~~The analyses of the association of the genetic variants with iron and ferritin in the GIS consortium~~ The analyses were adjusted for sex, age, and study-specific covariates such as genetic principal components to ~~adjust~~ control for population structure [13].

Since estimates of genetic effects from the discovery stage can be affected by an upward bias, referred to as "winner's curse" [23], sensitivity analyses were performed using estimates for the effects of the variants on iron and ferritin from the replication sample only, which included data on 24,986 individuals from 8 studies.

The strength of each instrument was evaluated using the F statistic, which is a function of the magnitude and precision of the genetic effect on the biomarker: $F=R^2(N-2)/(1-R^2)$, where R^2 is the proportion of the variance of iron (or ferritin) explained by the genetic variant and N is the sample size of the gene-iron (or gene-ferritin) association [24]. The R^2 values were estimated using the formula $2\times EAF\times(1-EAF)\times\beta^2$, where EAF is the effect allele frequency and β is the estimated genetic effect on iron (or ferritin) [25].

Estimates for the gene-eGFR association

The effects of the variants on eGFR were obtained from the CKDGen consortium meta-analysis, which included 74,354 individuals from 26 population-based studies [14]. The genetic effects ~~and~~ refer to per-allele effects on log-transformed eGFR. eGFR was estimated based on serum creatinine using the Modification of Diet in Renal Disease formula [26]. Similarly to iron and ferritin, the genetic association analyses for eGFR were adjusted for sex,

age, and study-specific covariates such as genetic principal components to adjust for population structure [14]. Effects of the same variants on CKD, defined as eGFR<60 ml/min/1.73 m^2 and corresponding to CKD stage 3-5, were obtained from the same meta-analysis, which included 6,271 CKD cases [14].

Mendelian Randomization analysis

MR estimates of the effect of iron on eGFR were derived for each instrument separately, using the Wald-type estimator, that is the ratio between the genetic effect on log(eGFR) (gene-eGFR association) and the genetic effect on serum iron levels (gene-iron association). The standard error of each MR estimate was derived using the Delta method [27]. The MR estimates were then combined using fixed-effect inverse-variance weighted meta-analysis. The same approach was used for the MR analysis of ferritin.

We also performed MR analyses to assess the effect of both iron and ferritin on the risk of CKD. However, these were only considered as secondary analyses because of the very low statistical power. The power calculation was performed using the online calculator for MR studies available at https://sb452.shinapps.io/power [28].

All analyses were performed using Stata 13 (StataCorp LP, College Station, TX, USA).

Results

The studies included in the GIS consortium for iron and ferritin had a mean age varying from 14 (SD: 5) to 68 (SD: 13) years. Detailed information on study-specific iron and ferritin levels by sex are reported in the supplement of reference [13], available at www.ncbi.nlm.nih.gov/pmc/articles/PMC4215164/bin/NIHMS619770-supplement-1.pdf (page 18). The studies included in the CKDGen consortium for eGFR and CKD were

characterized by a mean age ranging from 46 (SD: 16) to 76 (SD: 5) years. Mean eGFR varied from 96.5 (SD: 39.9) to 71.2 (SD: 14.8) ml/min/1.73m^2 and the median prevalence of CKD cases was 7.9% varying from 3.1% to 25.0% across studies. Detailed study specific information can be found in the supplementary table 3 of reference [14] (http://journals.plos.org/plosgenetics/article/asset?unique&id=info:doi/10.1371/journal.pgen.1002584.s015).

Of the five variants associated with iron at genome-wide significance level in the combined discovery and replication analysis of the GIS consortium (*TMPRSS6* rs855791; *HFE* rs1800562 and 1799945; *TF* rs8177240; *TFR2* rs7385804), the rs7385804 variant was excluded due to likely pleiotropic effects of the *TFR2* on kidney function independent from iron. This variant is reported as a *cis*-eQTL in blood and is associated with the expression of the *SLC12A9* and *EPHB4* genes. In animal studies these two genes have been shown to increase blood pressure, therefore potentially affecting glomerular filtration rate (*SLC12A9*) [29, 30], and to promote survival of podocytes in glomerular disease models (*EPHB4*) [31]. For the other four instruments, we found no evidence of pleiotropic effects, based on current biological knowledge from published literature and bioinformatics databases. Detailed findings for all variants are reported in the **Supplementary material** online.

All four genetic variants represent very strong instruments, with F statistics ranging between 98 and 1,000 (**Table 1**); these values are much higher than the conventional threshold of 10 used as a rule of thumb to distinguish between strong as weak instruments [32]. The four variants jointly explain 4% of the variance of iron levels in the GIS consortium [13]. **Table 1** reports the estimates of their per-allele effect on SD units of iron from the combined discovery and replication analyses of the GIS consortium (48,978 individuals), and per-allele effect on (log)eGFR from the discovery analysis of the CKDGen consortium (74,354 individuals) [14]. The MR estimate obtained from each variant and reported in **Table 1**

represents the effect of iron on log(eGFR), expressed as the change in log(eGFR) per SD unit increase in iron. The pooled MR estimate across the four variants (**Figure 1a**) showed a highly statistically significant protective effect of a SD unit of iron on log(eGFR) of 0.013 (95% CI: 0.004-0.021; *p-value*: 0.004), which translates in an eGFR increase of 1.3% (95% CI: 0.4-2.1%) per SD increase in iron. We observed high between-instrument heterogeneity (I^2: 67%; *p-het*: 0.028). The pooled MR estimate was driven by the two *HFE* variants, which showed highly consistent estimates; the MR estimates from the *TMPRSS6* and the *TF* variants showed no effect, although the latter had very large confidence intervals thus carrying little information. The sensitivity analysis using estimates only from the replication stage of the GIS consortium showed the same pooled MR estimate as the main analysis, 1.3% increase in eGFR per SD increase in iron (95% CI: 0.4-2.3%; *p-value*: 0.007).

All of the six variants considered for ferritin (*TMPRSS6* rs855791; *HFE* rs1800562 and rs1799945; *WDR75-SLC40A1* rs744653; *ABO* rs651007; *TEX14* rs411988) were used as instruments in our MR analysis since we found no evidence of pleiotropic effects for any of them (**Supplementary material** online). The estimates for the gene-ferritin association are reported in **Table 2,** and they refer to per-allele effects on SD units of $\log_{10}$-transformed ferritin from the combined discovery and replication analysis in the GIS consortium. Similarly to iron, the MR analysis showed a protective and statistically significant effect of ferritin on log(eGFR) (**Figure 1b**). The pooled MR estimate of the effect of a SD unit of $\log_{10}$(ferritin) on log(eGFR) was 0.021 (95% CI: 0.004-0.037; *p-value*: 0.014). As for iron, there was evidence of heterogeneity across MR estimates (I^2: 71%; *p-het*: 0.004). The results were the same in the sensitivity analysis performed using estimates from the replication stage only, with a pooled MR estimate of 0.023 (95% CI: 0.004-0.041; *p-value*: 0.018).

The results of our secondary analyses evaluating the effects of iron and ferritin on the risk of CKD were not statistically significant (**Supplementary Tables 1** and **2**). The odds ratio (OR) of CKD for a SD unit increase in iron was 1.01 (95% CI: 0.87-1.16) (**Supplementary Figure 1a**). For ferritin, the OR was 0.91 (95% CI: 0.70-1.20) per SD unit increase in $\log_{10}$(ferritin) (**Supplementary Figure 1b**). Such non-significant results are not surprising given the low statistical power of these analyses. Despite the high number of prevalent cases included in the analyses, the power of the MR approach to detect a 10% decrease in CKD risk per unit SD increase in iron, at a significance level of 5%, was of only 28%.

Discussion

Our study suggests a causal protective effect of circulating and stored iron on kidney function. These findings contrast with the observation that iron overload, resulting in hemosiderosis, can damage the kidney [33] through generation of reactive oxygen species in proximal tubular cells [3] and modulation of immune effector function of circulating monocytes [34]. They are, however, in line with evidence showing that also inadequate levels of iron can cause oxidative stress and, in particular, mitochondrial dysfunction due to mitochondrial DNA damage associated with oxidative stress [7, 8]. Both iron ~~deficiency~~ depletion and iron excess might therefore be detrimental to tissues. Increased cellular iron availability might be necessary for mitochondrial proteins, which contain iron/sulphur clusters as a cofactor. If accumulated above normal cellular levels, though, mitochondrial superoxide produced by the respiratory chain could result in the release of toxic soluble iron from iron/sulphur clusters, which is an additional source of reactive oxidative species generation [35].

Although very promising, our findings require confirmation. Despite our efforts to minimise the possibility of bias and in particular to exclude pleiotropic effects, we did observe high

between-instrument heterogeneity. Our thorough review of known biological effects did not find evidence of pleiotropy for any of the genetic variants used as instruments, but we cannot rule out the possibility of currently unknown pleiotropic effects. Knowledge of the biological function of genes is rapidly evolving but still largely incomplete, as is the knowledge of the effects of a variant on the expression of other genes (eQTL) within specific tissues, such as kidney. Moreover, pleiotropy is not the only possible explanation for the observed between-instrument heterogeneity, which could be due to other causes of departures from the assumptions underlying MR, including canalization, gene–environment interactions and population stratification [10]. A follow up of the findings of our study in experimental and functional models would provide conclusive evidence and allow understanding of the underlying mechanism of action of iron on kidney function.

It should also be noted that, although highly statistically significant, the effects of iron and ferritin on eGFR reported in our study are of limited magnitude. Further MR studies covering populations with wider ranges of eGFR and serum iron levels could better assess the dose-response relationship of the effect of iron on kidney function and the possible presence of threshold effects, e.g. a level of iron above which its effect on eGFR becomes minimal or even detrimental. This will require studies based on sample sizes even larger than those used here, or based on study populations with a larger number of CKD cases and of subjects with ~~iron deficiency or overload~~ altered iron levels. In general, although Mendelian randomization is a valuable tool for assessing causality when randomized clinical trials are not an option, the magnitude of the causal effect estimate that it provides may not correspond to the effect of an intervention modifying the exposure of interest [36]. Our Mendelian randomization study estimates the lifetime causal effect of serum iron levels on kidney function in the general population, an effect which could be very different from the effect, for example, of acute

changes in iron levels induced by iron supplementation in a population at risk of developing CKD.

~~Another~~ A possible limitation of our study is that data available from the CKDGen, which is the largest GWA consortium on kidney function, were based on eGFR estimated using the MDRD-4 equation. A different method has been recently proposed, the CKD-EPI equation, which has shown superior accuracy in subjects' classification. Although the use of GWA data for eGFR based on CKD-EPI could provide more accurate MR estimates, we would not expect differences in the direction of the effect, given the high degree of concordance between MDRD-4 and CKD-EPI estimated GFR [~~36~~37].

Confirmation of our findings ~~could have very important~~ has practical importance at both clinical and public health ~~implications~~levels. At a clinical level, establishing the exact role of iron metabolism in the regulation of kidney function is crucial given that iron currently tends to be considered detrimental, with some authors suggesting the use of iron chelators for the prevention and treatment of kidney disease [~~37~~38]. Iron metabolism in the kidney is characterised by complex processes and there may be a narrow window for "optimal iron levels", which could explain the apparently contradictory findings on whether iron is protective or detrimental for kidney function. Conclusive evidence on a protective role of iron on kidney function could also impact population health policies, since both absolute and functional iron deficiency is the most common nutritional deficiency in the world, particularly in women and children, and simple iron supplementation could easily address the issue [~~38~~39].

In conclusion, our findings suggest a beneficial effect of higher circulating and stored iron levels on kidney function in the general population. ~~Confirmation of these findings by experimental and functional studies and investigation of the underlying mechanism of action could have important clinical and public health implications.~~ Further research is required to confirm this causal association, investigate it in study populations at higher risk of CKD, and explore its underlying mechanism of action.

ACKNOWLEDGEMENTS

Collaboration between EURAC and the University of Leicester was supported by a travel grant from the Royal Society. Beben Benyamin is supported by the Australian National Health and Medical Research Council (NHMRC) and Motor Neuron Disease Research Australia (MNDRIA).

DISCLOSURE

Kathrin Eller received honoraria for speaking from Novartis and Amgen. All the remaining authors declared no competing financial interest.

References

1. Coyne DW, Kapoian T, Suki W*et al.* Ferric gluconate is highly efficacious in anemic hemodialysis patients with high serum ferritin and low transferrin saturation: results of the Dialysis Patients' Response to IV Iron with Elevated Ferritin (DRIVE) Study. *J Am Soc Nephrol* 2007; 18: 975–984

2. Stancu S, Bârsan L, Stanciu A, Mircescu G. Can the response to iron therapy be predicted in anemic nondialysis patients with chronic kidney disease? *Clin J Am Soc Nephrol* 2010; 5: 409–416

3. Sponsel HT, Alfrey AC, Hammond WS, Durr JA, Ray C, Anderson RJ. Effect of iron on renal tubular epithelial cells. *Kidney Int* 1996; 50: 436–444

4. Nankivell BJ, Chen J, Boadle RA, Harris DC. The role of tubular iron accumulation in the remnant kidney. *J Am Soc Nephrol* 1994; 4: 1598–1607

5. Hingorani S, Molitoris BA, Himmelfarb J. Ironing out the pathogenesis of acute kidney injury. *Am J Kidney Dis* 2009; 53: 569–571

6. Shah SV, Rajapurkar MM, Baliga R. The role of catalytic iron in acute kidney injury. *Clin J Am Soc Nephrol* 2011; 6: 2329–2331

7. Walter PB, Knutson MD, Paler-Martinez A *et al.* Iron deficiency and iron excess damage mitochondria and mitochondrial DNA in rats. *Proc Natl Acad Sci U S A* 2002; 99: 2264–2269

8. Knutson MD, Walter PB, Ames BN, Viteri FE. Both iron deficiency and daily iron supplements increase lipid peroxidation in rats. *J Nutr* 2000; 130: 621–628

9. Rensvold JW, Ong SE, Jeevananthan A, Carr SA, Mootha VK, Pagliarini DJ. Complementary RNA and protein profiling identifies iron as a key regulator of mitochondrial biogenesis. *Cell rep* 2013; 3: 237–245

10. Davey Smith G, Ebrahim S. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; 32: 1–22

11. Sheehan NA, Didelez V, Burton PR, Tobin MD. Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med* 2008; 5(8):e177

12. Burgess S, Thompson SG. Bias in causal estimation from Mendelian randomization studies with weak instruments. *Stat Med* 2011; 30: 1312–1323

13. 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, DOI: 10.1038/ncomms5926

14. Pattaro C, Klöttgen A, Teumer A *et al.* Genome-Wide Association and Functional Follow-Up Reveals New Loci for Kidney Function. *PLoS Genet* 2012; 8:e1002584. doi:10.1371/journal.pgen.1002584

15. Waheed A, Parkkila S, Zhou XY *et al.* Hereditary hemochromatosis: effects of C282Y and H63D mutations on association with beta2-microglobulin, intracellular processing, and cell surface expression of the HFE protein in COS-7 cells. *Proc Natl Acad Sci U S A* 1997; 94: 12384–12389

16. Feder JN, Tsuchihashi Z, Irrinki A *et al.* The hemochromatosis founder mutation in HLA-H disrupts beta2-microglobulin interaction and cell surface expression. *J Biol Chem* 1997; 272: 14025–14028

17. Feder JN, Gnirke A, Thomas W *et al.* A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 1996; 13: 399–408

18. Finberg KE, Heeney MM, Campagna DR, *et al.* Mutations in TMPRSS6 cause iron-refractory iron deficiency anemia (IRIDA). *Nat Genet* 2008; 40: 569–571

19. http://www.gtexportal.org/home/

20. Westra HJ, Peters MJ, Esko T *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013; 45: 1238–1243

21. https://genome.ucsc.edu/ENCODE/

22. Del Greco-M F, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med* 2015; 34: 2926–2940

23. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001; 29: 306–309

24. Palmer TM, Debbie A Lawlor DA, Roger M Harbord RM *et al.* Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* 2012; 21: 223–242

25. Falconer DS, Mackay TFC. Introduction to Quantitative Genetics 4th ed. Longman: Harlow, Essex, UK, 1996.

26. Levey AS, Bosh JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med* 1999; 130: 461–470

27. Thomas D, Conti D. Commentary: the concept of Mendelian randomization. *Intern J Epidem* 2004; 33: 21–25

28. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and binary outcome. *Intern J Epidem* 2014; 43: 922–929

29. Meyer JW, Flagella M, Sutliff RL *et al.* Decreased blood pressure and vascular smooth muscle tone in mice lacking basolateral Na(+)-K(+)-2Cl(-) cotransporter. *Am J Physiol Heart Circ Physiol* 2002; 283: H1846–1855

30. Orlov SN, Koltsova SV, Kapilevich LV, Gusakova SV, Dulin NO. NKCC1 an

NKCC2: The pathogenetic role of cation-cloride cotransporters in hypertension. *Genes Dis* 2015; 2: 186-196

31. Wnuk M, Hlushchuk R, Janot M *et al.* Podocyte EphB4 signaling helps recovery from glomerular injury. *Kidney Int* 2012; 81: 1212–1225

32. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008; 27: 1133–1163

33. Rostoker G, Griuncelli M, Loridon C *et al.* Hemodialysis-associated hemosiderosis in the era of erythropoiesis-stimulating agents: a MRI study. *Am J Med* 2012; 125: 991–999.e1

34. Sonnweber T, Theurl I, Seifert M *et al.* Impact of iron treatment on immune effector function and cellular iron status of circulating monocytes in dialysis patients. *Nephrol Dial Transplant* 2011; 26: 977–987

35. Rinnerthaler M, Hartl J, Stincone A *et al.* Mitochondria in ageing: there is metabolism beyond the ROS. FEMS *Yeast Res* 2013; doi:10.1111/1567-1364.12134

36. Burgess S, Thompson SG. Mendelian Randomization: Methods for using genetic variants in causal estimation. Chapman and Hall/CRC; 1 edition (6 Mar. 2015)

~~36~~37. Pattaro C, Stifter G, Modenese M, Minelli C, Pramstaller PP. Estimating the glomerular filtration rate in the general population using different equations: effects on classification and association. *Nephron Clin Pract* 2013; 123: 102-111

~~37~~38. Shah SV, Rajapurkar MM. The role of labile iron in kidney disease and treatment with chelation. *Hemoglobin* 2009; 33: 378–385

~~38~~39. Viteri FE. Iron supplementation for the control of iron deficiency in populations at risk. *Nutr Rev* 1997; 55: 195–209

Table 1. MR estimates of the *iron-eGFR* association (effect of a SD unit increase in iron on log(eGFR)) for the four variants, together with estimates of the *gene-iron* association (per-allele effect on SD units of iron) and the *gene-eGFR* association (per-allele effect on log(eGFR)).

Gene SNP	EA	EAF	gene-iron (N=48,978)				gene-eGFR (N=74,354)		MR iron-eGFR	
			R^2	F	Beta (SE)	*p*	Beta (SE)	*p*	Beta (SE)	*p*
TMPRSS6 *r*s855791	G	0.55	0.02	1,000	0.181 (0.007)	1.32×10^{-139}	0.0002 (0.001)	0.849	0.001 (0.007)	0.878
HFE rs1800562	A	0.07	0.01	495	0.328 (0.016)	2.72×10^{-97}	0.007 (0.003)	0.004	0.023 (0.008)	0.005
HFE rs1799945	G	0.15	0.01	495	0.189 (0.010)	1.10×10^{-81}	0.005 (0.002)	0.003	0.026 (0.009)	0.004
TF rs8177240	G	0.33	0.002	98	0.066 (0.007)	6.65×10^{-20}	-0.001 (0.001)	0.304	-0.020 (0.020)	0.320

Abbreviations: SNP: single nucleotide polymorphism; EA: effect allele (allele increasing iron levels); EAF: effect allele frequency; Beta (SE): estimate (standard error); *p*: p-value; R^2: percentage of the variance of iron explained by the SNP; F: F statistic.

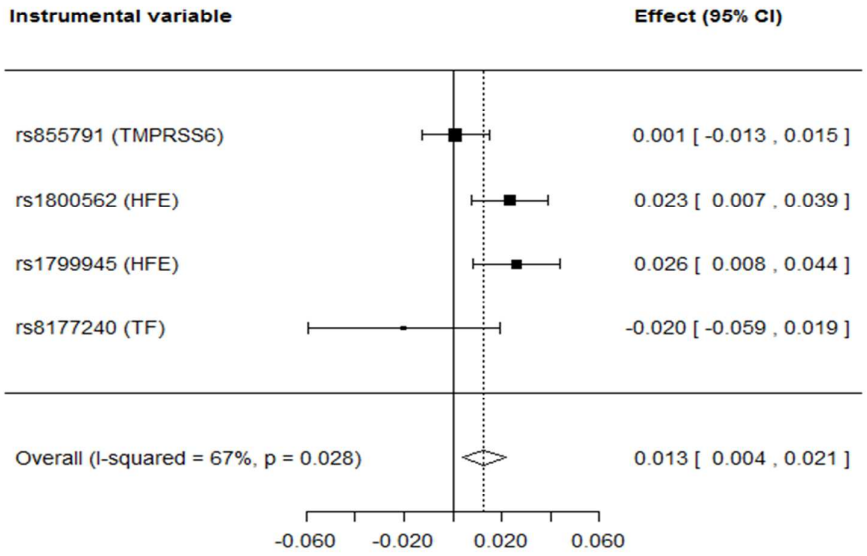
Table 2. MR estimates of the *ferritin-eGFR* association (effect of a SD unit increase in $\log_{10}$(ferritin) on log(eGFR)) for the six variants, together with estimates of the *gene-ferritin* association (per-allele effect on SD units of $\log_{10}$(ferritin)) and the *gene-eGFR* association (per-allele effect on log(eGFR)).

Gene SNP	EA	EAF	gene-ferritin (N=48,978)				gene-eGFR (N=74,354)		MR ferritin-eGFR	
			R^2	F	Beta (SE)	*p*	Beta (SE)	*p*	Beta (SE)	*p*
HFE rs1800562	A	0.07	0.010	495	0.204 (0.016)	1.54×10^{-38}	0.007 (0.003)	0.004	0.036 (0.013)	0.005
WDR75-SLC40A1 rs744653	C	0.15	0.002	98	0.089 (0.010)	8.37×10^{-19}	-0.0021 (0.0017)	0.221	-0.024 (0.019)	0.221
TMPRSS6 rs855791	G	0.55	0.001	49	0.055 (0.007)	1.38×10^{-14}	0.0002 (0.001)	0.849	0.004 (0.024)	0.878
TEX14 rs411988	G	0.44	0.001	49	0.044 (0.007)	1.59×10^{-10}	0.003 (0.001)	0.012	0.068 (0.029)	0.020
HFE rs1799945	G	0.15	0.001	49	0.065 (0.01)	1.71×10^{-10}	0.005 (0.002)	0.004	0.077 (0.029)	0.007
ABO rs651007	C	0.80	0.001	49	0.05 (0.009)	1.31×10^{-8}	-0.002 (0.002)	0.205	-0.038 (0.031)	0.217

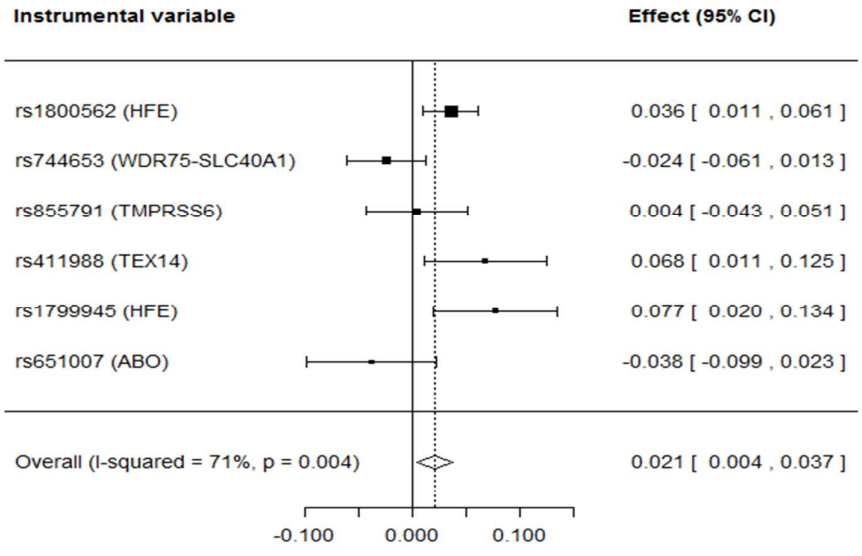
Abbreviations: SNP: single nucleotide polymorphism; EA: effect allele (allele increasing ferritin levels); EA: effect allele frequency; Beta (SE): estimate (standard error) of the association; *p*: p-value; R^2: percentage of the variance of the ferritin explained by the SNP; F: F statistic.

Figure 1: Forest plot for the meta-analysis of MR estimates across variants for: a) *iron-eGFR*; b) *ferritin-eGFR*. The size of the squares is proportional to the precision of the estimate, and the horizontal line indicates its 95%CI; the centre of the diamond represents the pooled MR estimate and the lateral tips indicate its 95% CI.

a) Effect of a SD increase in iron on log(eGFR)



b) Effect of a SD increase in $\log_{10}$(ferritin) on log(eGFR)



Serum Iron Level and Kidney Function: a Mendelian Randomization Study

Fabiola Del Greco M[1*] PhD, Luisa Foco[1] PhD, Irene Pichler[1] PhD, Philipp Eller[2] MBA, Kathrin Eller[2] MD, Beben Benyamin[3,4]PhD, John B Whitfield[4] PhD, Genetics of Iron Status Consortium, CKDGen Consortium, Peter P Pramstaller[1] MD, John R Thompson[5] PhD, Cristian Pattaro[1] PhD, Cosetta Minelli[6*] PhD

1. Center for Biomedicine, European Academy of Bolzano/Bozen (EURAC), affiliated to the University of Lübeck, Bolzano, Italy
2. Medical University of Graz, Graz, Austria
3. Queensland Brain Institute, The University of Queensland, Queensland, Australia
4. QIMR Berghofer Medical Research Institute, Brisbane, Australia
5. Department of Health Sciences, University of Leicester, Leicester, UK
6. Respiratory Epidemiology, Occupational Medicine and Public Health, National Heart and Lung Institute, Imperial College, London, United Kingdom

Corresponding authors:

Fabiola Del Greco M, PhD
Center for Biomedicine – EURAC research Bozen/Bolzano
Via Galvani 31
39100 Bolzano, Italy
E-mail: fabiola.delgreco@eurac.edu
Tel: +39 0471 055 528; Fax: +39 0471 055 599

Cosetta Minelli, MD PhD
NHLI, Imperial College London
Emmanuel Kaye Building
1 Manresa Road
SW3 6LR London, United Kingdom
E-mail: cosetta.minelli1@imperial.ac.uk
Tel: +44 (0) 207 5947758

Abstract word count: 244 words

Total word count: 3,492 words

Running title: Iron and eGFR

Keywords: Mendelian randomization; iron; ferritin; glomerular filtration rate; kidney function; chronic kidney disease

Abstract

Background. Iron depletion is a known consequence of chronic kidney disease (CKD), but there is contradicting epidemiological evidence on whether iron itself affects kidney function and whether its effect is protective or detrimental in the general population. While epidemiological studies tend to be affected by confounding and reverse causation, Mendelian Randomization (MR) can provide unconfounded estimates of causal effects by using genes as instruments.

Methods. We performed a MR study of the effect of serum iron levels on estimated glomerular filtration rate (eGFR), using genetic variants known to be associated with iron. MR estimates of the effect of iron on eGFR were derived based on the association of each variant with iron and eGFR from two large genome-wide meta-analyses on 48,978 and 74,354 individuals, respectively. We performed a similar MR analysis for ferritin, which measures iron stored in the body, using variants associated with ferritin.

Results. A combined MR estimate across all variants showed a 1.3% increase in eGFR per standard deviation increase in iron (95%CI: 0.4-2.1%; *p-value*: 0.004). The results for ferritin were consistent with those for iron. Secondary MR analyses of the effects of iron and ferritin on CKD did not show significant associations but had very low statistical power.

Conclusions. Our study suggests a protective effect of iron on kidney function in the general population. Further research is required to confirm this causal association, investigate it in study populations at higher risk of CKD, and explore its underlying mechanism of action.

Introduction

It is well known that chronic kidney disease (CKD) can cause iron depletion, and iron supplementation is widely used to prevent and treat anaemia in CKD patients [1, 2]. What is not known is whether iron can influence kidney function and disease in the general population, and it is unclear whether altered iron levels could also be a cause, rather than a consequence, of kidney disease. The deleterious effects of iron accumulation in tissues in iron-overload conditions is well recognized, and experimental models of both acute kidney injury and CKD have suggested that iron could be nephrotoxic even in the absence of systemic overload, possibly through the formation of free radicals [3-6]. However, studies in rats have shown that dietary iron elimination can cause tissue damage through oxidative stress and mitochondrial damage [7, 8]. Cellular iron homeostasis plays an important role in mitochondrial biogenesis, and iron deprived cells show downregulation of mitochondrial protein levels and oxidative capacity [9].

Assessing the effects of iron on kidney function in humans using observational data is very challenging because of the difficulty in disentangling causal from spurious effects due to confounding and reverse causation. On the other hand, randomized clinical trials investigating the effect of iron supplementation on kidney function in healthy individuals would be difficult to justify. In situations like this, Mendelian randomization (MR) is being increasingly used to provide estimates of causal effects of modifiable risk factors (e.g. biomarkers) on disease traits. MR is an instrumental variable approach based on the use of genes as 'proxies' (instruments) for the biomarker: genetic variants known to be associated with altered levels of the biomarker are investigated for their effect on the clinical endpoint [10, 11]. Being carried from birth, genes cannot be influenced by changes in the clinical endpoint (no reverse causation) or life-style and environmental predictors (no classical confounding). Moreover,

the MR approach estimates the lifetime effect of the risk factor on the clinical endpoint, as opposed to the effect at a certain point in life estimated by classical observational studies.

The rapidly accumulating evidence on genetic associations for biomarkers (and other risk factors) and disease traits explains the increasing interest in the use of MR to distinguish between spurious and causal associations. As any other instrumental variable approach, MR suffers from limited statistical power [10], but the availability of very large genetic datasets currently ensures adequate power to investigate the causal effects on disease of many risk factors. However, the validity of MR heavily depends on the validity of the instruments chosen, which makes the selection of the genetic variants the most important step of a MR study. A key aspect is the "strength" of the instrument (the genetic variant), defined as a function of both the magnitude of the genetic effect on the risk factor and the precision with which it is estimated; the use of weak instruments can introduce bias (weak-instrument bias) [12]. Given the modest genetic effects typical of common variants, adequate instrument strength is usually achieved through large sample sizes for the gene-risk factor association. Another fundamental aspect is the absence of pleiotropy, which means that the instrument influences the disease trait only through the risk factor considered and not through any other pathway. This is the most challenging assumption in MR and its violation can produce biased MR estimates [10, 11].

In this study we have used MR to investigate the lifetime causal effect of serum iron levels on kidney function in the general population. Kidney function was evaluated based on the glomerular filtration rate, estimated from serum creatinine using the four-parameter MDRD equation (eGFR). We used genetic data from the largest genome-wide association (GWA) consortia for iron markers (GIS consortium) [13] and eGFR (CKDGen consortium) [14]. As

instruments we selected all variants found associated with serum iron levels and replicated in the GIS consortium, with the exception of one variant which we excluded due to possible pleiotropic effects on kidney. To investigate whether findings for the effect of circulating iron levels on kidney function could be also observed for iron stored in the body, we performed a similar MR analysis for the effect of serum ferritin levels on eGFR, again using all genetic variants associated with ferritin and replicated in the GIS consortium. Finally, we repeated the MR analyses to estimate the effects of both iron and ferritin on the risk of CKD, defined as eGFR<60 ml/min/1.73 m^2.

Methods

Selection of genetic variants and investigation of pleiotropy

We considered as instruments for the MR analyses all genetic variants reaching genome-wide significance ($p<5x10^{-8}$) in the combined discovery and replication analysis of the GIS consortium on serum iron and ferritin, based on a total sample size of 48,978 individuals from 18 studies, which represents the largest GWA meta-analysis conducted so far on these phenotypes [13].

For the MR analysis of the effect of iron on eGFR, five variants were considered: rs855791 (also known as V736A) in the transmembrane protease 6 (*TMPRSS6*) gene; rs1800562 (C282Y) and rs1799945 (H63D) in the hemochromatosis (*HFE*) gene; rs8177240 in the transferrin (*TF*) gene; rs7385804 in the transferrin receptor 2 (*TFR2*) gene. All variants are independent, that is they are not in linkage disequilibrium (LD), including the two variants in *HFE* (LD $r^2=0.01$) located in genomic regions coding for functionally distinct protein domains [15, 16]. These variants were associated with serum iron levels at p-values ranging

from 10^{-20} to 10^{-139}. Three of them, rs855791, rs1800562, and rs1799945 were already known to affect serum iron levels [17, 18].

For the MR analysis of the effect of ferritin on eGFR, six variants were considered, three of which were the same as for iron, *TMPRSS6* rs855791, *HFE* rs1800562, and *HFE* rs1799945; the other three included a variant located between the WD Repeat Domain 75 (*WDR75*) and the Solute Carrier Family 40 (Iron-Regulated Transporter) Member 1 (*SLC40A1*) genes, rs744653, a variant in the ABO Blood Group (*ABO*) gene, rs651007, and a variant in the Testis Expressed 14 (*TEX14*) gene, rs411988. These six variants were associated with ferritin with p-values ranging from 10^{-8} to 10^{-38}.

For all variants for iron and ferritin, we investigated the possible presence of pleiotropy both *a priori* based on biological knowledge, and *a posteriori* based on statistical evidence. For the biological assessment, we performed a literature search to identify evidence on the effects of the genes in which the selected variants are annotated. In cases where the variants were acting as *cis* and/or *trans* expression quantitative trait loci (eQTLs), we also included in our investigation any additional gene showing differential expression. For the latter task, we used several freely available online tools, including GTEx [19], Blood eQTL database [20], and ENCODE project data at UCSC Genome browser [21]. For the statistical assessment of possible unknown pleiotropic effects, or other factors invalidating the MR assumption that any effect of the variant on eGFR is mediated by iron (or ferritin), we investigated the heterogeneity of MR estimates across instruments, using the heterogeneity Q test (*p-het*) and the I^2 index as we have previously described [22].

Estimates for the gene-iron and gene-ferritin associations

The effects of the genetic variants on serum iron and ferritin were based on the GIS GWA meta-analysis on 48,978 individuals from the 18 studies. The analyses were adjusted for sex, age, and study-specific covariates such as genetic principal components to control for population structure [13].

Since estimates of genetic effects from the discovery stage can be affected by an upward bias, referred to as "winner's curse" [23], sensitivity analyses were performed using estimates for the effects of the variants on iron and ferritin from the replication sample only, which included data on 24,986 individuals from 8 studies.

The strength of each instrument was evaluated using the F statistic, which is a function of the magnitude and precision of the genetic effect on the biomarker: $F=R^2(N-2)/(1-R^2)$, where R^2 is the proportion of the variance of iron (or ferritin) explained by the genetic variant and N is the sample size of the gene-iron (or gene-ferritin) association [24]. The R^2 values were estimated using the formula $2\times EAF\times(1-EAF)\times\beta^2$, where EAF is the effect allele frequency and β is the estimated genetic effect on iron (or ferritin) [25].

Estimates for the gene-eGFR association

The effects of the variants on eGFR were obtained from the CKDGen consortium meta-analysis, which included 74,354 individuals from 26 population-based studies [14]. The genetic effects refer to per-allele effects on log-transformed eGFR. eGFR was estimated based on serum creatinine using the Modification of Diet in Renal Disease formula [26]. Similarly to iron and ferritin, the genetic association analyses for eGFR were adjusted for sex, age, and study-specific covariates such as genetic principal components to adjust for population

structure [14]. Effects of the same variants on CKD, defined as eGFR<60 ml/min/1.73 m^2 and corresponding to CKD stage 3-5, were obtained from the same meta-analysis, which included 6,271 CKD cases [14].

Mendelian Randomization analysis

MR estimates of the effect of iron on eGFR were derived for each instrument separately, using the Wald-type estimator, that is the ratio between the genetic effect on log(eGFR) (gene-eGFR association) and the genetic effect on serum iron levels (gene-iron association). The standard error of each MR estimate was derived using the Delta method [27]. The MR estimates were then combined using fixed-effect inverse-variance weighted meta-analysis. The same approach was used for the MR analysis of ferritin.

We also performed MR analyses to assess the effect of both iron and ferritin on the risk of CKD. However, these were only considered as secondary analyses because of the very low statistical power. The power calculation was performed using the online calculator for MR studies available at https://sb452.shinapps.io/power [28].

All analyses were performed using Stata 13 (StataCorp LP, College Station, TX, USA).

Results

The studies included in the GIS consortium for iron and ferritin had a mean age varying from 14 (SD: 5) to 68 (SD: 13) years. Detailed information on study-specific iron and ferritin levels by sex are reported in the supplement of reference [13], available at www.ncbi.nlm.nih.gov/pmc/articles/PMC4215164/bin/NIHMS619770-supplement-1.pdf (page 18). The studies included in the CKDGen consortium for eGFR and CKD were characterized by a mean age ranging from 46 (SD: 16) to 76 (SD: 5) years. Mean eGFR varied from 96.5 (SD:

39.9) to 71.2 (SD: 14.8) ml/min/1.73m^2 and the median prevalence of CKD cases was 7.9% varying from 3.1% to 25.0% across studies. Detailed study specific information can be found in the supplementary table 3 of reference [14] (http://journals.plos.org/plosgenetics/article/asset?unique&id=info:doi/10.1371/journal.pgen.1002584.s015).

Of the five variants associated with iron at genome-wide significance level in the combined discovery and replication analysis of the GIS consortium (*TMPRSS6* rs855791; *HFE* rs1800562 and 1799945; *TF* rs8177240; *TFR2* rs7385804), the rs7385804 variant was excluded due to likely pleiotropic effects of the *TFR2* on kidney function independent from iron. This variant is reported as a *cis*-eQTL in blood and is associated with the expression of the *SLC12A9* and *EPHB4* genes. In animal studies these two genes have been shown to increase blood pressure, therefore potentially affecting glomerular filtration rate (*SLC12A9*) [29, 30], and to promote survival of podocytes in glomerular disease models (*EPHB4*) [31]. For the other four instruments, we found no evidence of pleiotropic effects, based on current biological knowledge from published literature and bioinformatics databases. Detailed findings for all variants are reported in the **Supplementary material** online.

All four genetic variants represent very strong instruments, with F statistics ranging between 98 and 1,000 (**Table 1**); these values are much higher than the conventional threshold of 10 used as a rule of thumb to distinguish between strong as weak instruments [32]. The four variants jointly explain 4% of the variance of iron levels in the GIS consortium [13]. **Table 1** reports the estimates of their per-allele effect on SD units of iron from the combined discovery and replication analyses of the GIS consortium (48,978 individuals), and per-allele effect on (log)eGFR from the discovery analysis of the CKDGen consortium (74,354 individuals) [14]. The MR estimate obtained from each variant and reported in **Table 1** represents the effect of iron on log(eGFR), expressed as the change in log(eGFR) per SD unit

increase in iron. The pooled MR estimate across the four variants (**Figure 1a**) showed a highly statistically significant protective effect of a SD unit of iron on log(eGFR) of 0.013 (95% CI: 0.004-0.021; *p-value*: 0.004), which translates in an eGFR increase of 1.3% (95% CI: 0.4-2.1%) per SD increase in iron. We observed high between-instrument heterogeneity (I^2: 67%; *p-het*: 0.028). The pooled MR estimate was driven by the two *HFE* variants, which showed highly consistent estimates; the MR estimates from the *TMPRSS6* and the *TF* variants showed no effect, although the latter had very large confidence intervals thus carrying little information. The sensitivity analysis using estimates only from the replication stage of the GIS consortium showed the same pooled MR estimate as the main analysis, 1.3% increase in eGFR per SD increase in iron (95% CI: 0.4-2.3%; *p-value*: 0.007).

All of the six variants considered for ferritin (*TMPRSS6* rs855791; *HFE* rs1800562 and rs1799945; *WDR75-SLC40A1* rs744653; *ABO* rs651007; *TEX14* rs411988) were used as instruments in our MR analysis since we found no evidence of pleiotropic effects for any of them (**Supplementary material** online). The estimates for the gene-ferritin association are reported in **Table 2,** and they refer to per-allele effects on SD units of $\log_{10}$-transformed ferritin from the combined discovery and replication analysis in the GIS consortium. Similarly to iron, the MR analysis showed a protective and statistically significant effect of ferritin on log(eGFR) (**Figure 1b**). The pooled MR estimate of the effect of a SD unit of $\log_{10}$(ferritin) on log(eGFR) was 0.021 (95% CI: 0.004-0.037; *p-value*: 0.014). As for iron, there was evidence of heterogeneity across MR estimates (I^2: 71%; *p-het*: 0.004). The results were the same in the sensitivity analysis performed using estimates from the replication stage only, with a pooled MR estimate of 0.023 (95% CI: 0.004-0.041; *p-value*: 0.018).

The results of our secondary analyses evaluating the effects of iron and ferritin on the risk of CKD were not statistically significant (**Supplementary Tables 1** and **2**). The odds ratio (OR) of CKD for a SD unit increase in iron was 1.01 (95% CI: 0.87-1.16) (**Supplementary Figure 1a**). For ferritin, the OR was 0.91 (95% CI: 0.70-1.20) per SD unit increase in $\log_{10}$(ferritin) (**Supplementary Figure 1b**). Such non-significant results are not surprising given the low statistical power of these analyses. Despite the high number of prevalent cases included in the analyses, the power of the MR approach to detect a 10% decrease in CKD risk per unit SD increase in iron, at a significance level of 5%, was of only 28%.

Discussion

Our study suggests a causal protective effect of circulating and stored iron on kidney function. These findings contrast with the observation that iron overload, resulting in hemosiderosis, can damage the kidney [33] through generation of reactive oxygen species in proximal tubular cells [3] and modulation of immune effector function of circulating monocytes [34]. They are, however, in line with evidence showing that also inadequate levels of iron can cause oxidative stress and, in particular, mitochondrial dysfunction due to mitochondrial DNA damage associated with oxidative stress [7, 8]. Both iron depletion and iron excess might therefore be detrimental to tissues. Increased cellular iron availability might be necessary for mitochondrial proteins, which contain iron/sulphur clusters as a cofactor. If accumulated above normal cellular levels, though, mitochondrial superoxide produced by the respiratory chain could result in the release of toxic soluble iron from iron/sulphur clusters, which is an additional source of reactive oxidative species generation [35].

Although very promising, our findings require confirmation. Despite our efforts to minimise the possibility of bias and in particular to exclude pleiotropic effects, we did observe high

between-instrument heterogeneity. Our thorough review of known biological effects did not find evidence of pleiotropy for any of the genetic variants used as instruments, but we cannot rule out the possibility of currently unknown pleiotropic effects. Knowledge of the biological function of genes is rapidly evolving but still largely incomplete, as is the knowledge of the effects of a variant on the expression of other genes (eQTL) within specific tissues, such as kidney. Moreover, pleiotropy is not the only possible explanation for the observed between-instrument heterogeneity, which could be due to other causes of departures from the assumptions underlying MR, including canalization, gene–environment interactions and population stratification [10]. A follow up of the findings of our study in experimental and functional models would provide conclusive evidence and allow understanding of the underlying mechanism of action of iron on kidney function.

It should also be noted that, although highly statistically significant, the effects of iron and ferritin on eGFR reported in our study are of limited magnitude. Further MR studies covering populations with wider ranges of eGFR and serum iron levels could better assess the dose-response relationship of the effect of iron on kidney function and the possible presence of threshold effects, e.g. a level of iron above which its effect on eGFR becomes minimal or even detrimental. This will require studies based on sample sizes even larger than those used here, or based on study populations with a larger number of CKD cases and of subjects with altered iron levels. In general, although Mendelian randomization is a valuable tool for assessing causality when randomized clinical trials are not an option, the magnitude of the causal effect estimate that it provides may not correspond to the effect of an intervention modifying the exposure of interest [36]. Our Mendelian randomization study estimates the lifetime causal effect of serum iron levels on kidney function in the general population, an

effect which could be very different from the effect, for example, of acute changes in iron levels induced by iron supplementation in a population at risk of developing CKD.

A possible limitation of our study is that data available from the CKDGen, which is the largest GWA consortium on kidney function, were based on eGFR estimated using the MDRD-4 equation. A different method has been recently proposed, the CKD-EPI equation, which has shown superior accuracy in subjects' classification. Although the use of GWA data for eGFR based on CKD-EPI could provide more accurate MR estimates, we would not expect differences in the direction of the effect, given the high degree of concordance between MDRD-4 and CKD-EPI estimated GFR [37].

Confirmation of our findings has practical importance at both clinical and public health levels. At a clinical level, establishing the exact role of iron metabolism in the regulation of kidney function is crucial given that iron currently tends to be considered detrimental, with some authors suggesting the use of iron chelators for the prevention and treatment of kidney disease [38]. Iron metabolism in the kidney is characterised by complex processes and there may be a narrow window for "optimal iron levels", which could explain the apparently contradictory findings on whether iron is protective or detrimental for kidney function. Conclusive evidence on a protective role of iron on kidney function could also impact population health policies, since both absolute and functional iron deficiency is the most common nutritional deficiency in the world, particularly in women and children, and simple iron supplementation could easily address the issue [39].

In conclusion, our findings suggest a beneficial effect of higher circulating and stored iron levels on kidney function in the general population. Further research is required to confirm

this causal association, investigate it in study populations at higher risk of CKD, and explore its underlying mechanism of action.

ACKNOWLEDGEMENTS

Collaboration between EURAC and the University of Leicester was supported by a travel grant from the Royal Society. Beben Benyamin is supported by the Australian National Health and Medical Research Council (NHMRC) and Motor Neuron Disease Research Australia (MNDRIA).

DISCLOSURE

Kathrin Eller received honoraria for speaking from Novartis and Amgen. All the remaining authors declared no competing financial interest.

References

1. Coyne DW, Kapoian T, Suki W*et al.* Ferric gluconate is highly efficacious in anemic hemodialysis patients with high serum ferritin and low transferrin saturation: results of the Dialysis Patients' Response to IV Iron with Elevated Ferritin (DRIVE) Study. *J Am Soc Nephrol* 2007; 18: 975–984

2. Stancu S, Bârsan L, Stanciu A, Mircescu G. Can the response to iron therapy be predicted in anemic nondialysis patients with chronic kidney disease? *Clin J Am Soc Nephrol* 2010; 5: 409–416

3. Sponsel HT, Alfrey AC, Hammond WS, Durr JA, Ray C, Anderson RJ. Effect of iron on renal tubular epithelial cells. *Kidney Int* 1996; 50: 436–444

4. Nankivell BJ, Chen J, Boadle RA, Harris DC. The role of tubular iron accumulation in the remnant kidney. *J Am Soc Nephrol* 1994; 4: 1598–1607

5. Hingorani S, Molitoris BA, Himmelfarb J. Ironing out the pathogenesis of acute kidney injury. *Am J Kidney Dis* 2009; 53: 569–571

6. Shah SV, Rajapurkar MM, Baliga R. The role of catalytic iron in acute kidney injury. *Clin J Am Soc Nephrol* 2011; 6: 2329–2331

7. Walter PB, Knutson MD, Paler-Martinez A *et al.* Iron deficiency and iron excess damage mitochondria and mitochondrial DNA in rats. *Proc Natl Acad Sci U S A* 2002; 99: 2264–2269

8. Knutson MD, Walter PB, Ames BN, Viteri FE. Both iron deficiency and daily iron supplements increase lipid peroxidation in rats. *J Nutr* 2000; 130: 621–628

9. Rensvold JW, Ong SE, Jeevananthan A, Carr SA, Mootha VK, Pagliarini DJ. Complementary RNA and protein profiling identifies iron as a key regulator of mitochondrial biogenesis. *Cell rep* 2013; 3: 237–245

10. Davey Smith G, Ebrahim S. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; 32: 1–22

11. Sheehan NA, Didelez V, Burton PR, Tobin MD. Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med* 2008; 5(8):e177

12. Burgess S, Thompson SG. Bias in causal estimation from Mendelian randomization studies with weak instruments. *Stat Med* 2011; 30: 1312–1323

13. 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, DOI: 10.1038/ncomms5926

14. Pattaro C, Klöttgen A, Teumer A *et al.* Genome-Wide Association and Functional Follow-Up Reveals New Loci for Kidney Function. *PLoS Genet* 2012; 8:e1002584. doi:10.1371/journal.pgen.1002584

15. Waheed A, Parkkila S, Zhou XY *et al.* Hereditary hemochromatosis: effects of C282Y and H63D mutations on association with beta2-microglobulin, intracellular processing, and cell surface expression of the HFE protein in COS-7 cells. *Proc Natl Acad Sci U S A* 1997; 94: 12384–12389

16. Feder JN, Tsuchihashi Z, Irrinki A *et al.* The hemochromatosis founder mutation in HLA-H disrupts beta2-microglobulin interaction and cell surface expression. *J Biol Chem* 1997; 272: 14025–14028

17. Feder JN, Gnirke A, Thomas W *et al.* A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 1996; 13: 399–408

18. Finberg KE, Heeney MM, Campagna DR, *et al.* Mutations in TMPRSS6 cause iron-refractory iron deficiency anemia (IRIDA). *Nat Genet* 2008; 40: 569–571

19. http://www.gtexportal.org/home/

20. Westra HJ, Peters MJ, Esko T *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013; 45: 1238–1243

21. https://genome.ucsc.edu/ENCODE/

22. Del Greco-M F, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med* 2015; 34: 2926–2940

23. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001; 29: 306–309

24. Palmer TM, Debbie A Lawlor DA, Roger M Harbord RM *et al.* Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* 2012; 21: 223–242

25. Falconer DS, Mackay TFC. Introduction to Quantitative Genetics 4th ed. Longman: Harlow, Essex, UK, 1996.

26. Levey AS, Bosh JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med* 1999; 130: 461–470

27. Thomas D, Conti D. Commentary: the concept of Mendelian randomization. *Intern J Epidem* 2004; 33: 21–25

28. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and binary outcome. *Intern J Epidem* 2014; 43: 922–929

29. Meyer JW, Flagella M, Sutliff RL *et al.* Decreased blood pressure and vascular smooth muscle tone in mice lacking basolateral Na(+)-K(+)-2Cl(-) cotransporter. *Am J Physiol Heart Circ Physiol* 2002; 283: H1846–1855

30. Orlov SN, Koltsova SV, Kapilevich LV, Gusakova SV, Dulin NO. NKCC1 an

NKCC2: The pathogenetic role of cation-cloride cotransporters in hypertension. *Genes Dis* 2015; 2: 186-196

31. Wnuk M, Hlushchuk R, Janot M *et al.* Podocyte EphB4 signaling helps recovery from glomerular injury. *Kidney Int* 2012; 81: 1212–1225

32. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008; 27: 1133–1163

33. Rostoker G, Griuncelli M, Loridon C *et al.* Hemodialysis-associated hemosiderosis in the era of erythropoiesis-stimulating agents: a MRI study. *Am J Med* 2012; 125: 991–999.e1

34. Sonnweber T, Theurl I, Seifert M *et al.* Impact of iron treatment on immune effector function and cellular iron status of circulating monocytes in dialysis patients. *Nephrol Dial Transplant* 2011; 26: 977–987

35. Rinnerthaler M, Hartl J, Stincone A *et al.* Mitochondria in ageing: there is metabolism beyond the ROS. FEMS *Yeast Res* 2013; doi:10.1111/1567-1364.12134

36. Burgess S, Thompson SG. Mendelian Randomization: Methods for using genetic variants in causal estimation. Chapman and Hall/CRC; 1 edition (6 Mar. 2015)

37. Pattaro C, Stifter G, Modenese M, Minelli C, Pramstaller PP. Estimating the glomerular filtration rate in the general population using different equations: effects on classification and association. *Nephron Clin Pract* 2013; 123: 102-111

38. Shah SV, Rajapurkar MM. The role of labile iron in kidney disease and treatment with chelation. *Hemoglobin* 2009; 33: 378–385

39. Viteri FE. Iron supplementation for the control of iron deficiency in populations at risk. *Nutr Rev* 1997; 55: 195–209

Table 1. MR estimates of the *iron-eGFR* association (effect of a SD unit increase in iron on log(eGFR)) for the four variants, together with estimates of the *gene-iron* association (per-allele effect on SD units of iron) and the *gene-eGFR* association (per-allele effect on log(eGFR)).

Gene SNP	EA	EAF	gene-iron (N=48,978)				gene-eGFR (N=74,354)		MR iron-eGFR	
			R^2	F	Beta (SE)	*p*	Beta (SE)	*p*	Beta (SE)	*p*
TMPRSS6 *r*s855791	G	0.55	0.02	1,000	0.181 (0.007)	1.32×10^{-139}	0.0002 (0.001)	0.849	0.001 (0.007)	0.878
HFE rs1800562	A	0.07	0.01	495	0.328 (0.016)	2.72×10^{-97}	0.007 (0.003)	0.004	0.023 (0.008)	0.005
HFE rs1799945	G	0.15	0.01	495	0.189 (0.010)	1.10×10^{-81}	0.005 (0.002)	0.003	0.026 (0.009)	0.004
TF rs8177240	G	0.33	0.002	98	0.066 (0.007)	6.65×10^{-20}	-0.001 (0.001)	0.304	-0.020 (0.020)	0.320

Abbreviations: SNP: single nucleotide polymorphism; EA: effect allele (allele increasing iron levels); EAF: effect allele frequency; Beta (SE): estimate (standard error); *p*: p-value; R^2: percentage of the variance of iron explained by the SNP; F: F statistic.

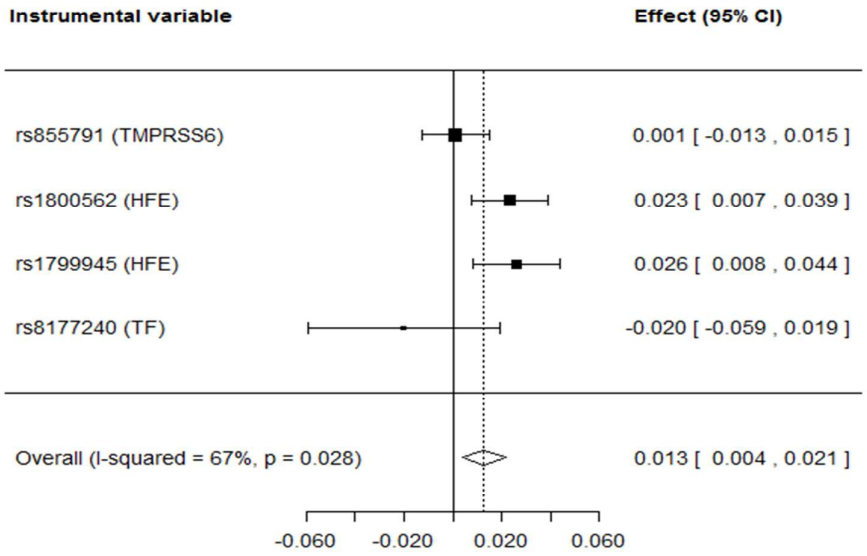
Table 2. MR estimates of the *ferritin-eGFR* association (effect of a SD unit increase in $\log_{10}$(ferritin) on log(eGFR)) for the six variants, together with estimates of the *gene-ferritin* association (per-allele effect on SD units of $\log_{10}$(ferritin)) and the *gene-eGFR* association (per-allele effect on log(eGFR)).

Gene SNP	EA	EAF	gene-ferritin (N=48,978)				gene-eGFR (N=74,354)		MR ferritin-eGFR	
			R^2	F	Beta (SE)	*p*	Beta (SE)	*p*	Beta (SE)	*p*
HFE rs1800562	A	0.07	0.010	495	0.204 (0.016)	1.54×10^{-38}	0.007 (0.003)	0.004	0.036 (0.013)	0.005
WDR75-SLC40A1 rs744653	C	0.15	0.002	98	0.089 (0.010)	8.37×10^{-19}	-0.0021 (0.0017)	0.221	-0.024 (0.019)	0.221
TMPRSS6 rs855791	G	0.55	0.001	49	0.055 (0.007)	1.38×10^{-14}	0.0002 (0.001)	0.849	0.004 (0.024)	0.878
TEX14 rs411988	G	0.44	0.001	49	0.044 (0.007)	1.59×10^{-10}	0.003 (0.001)	0.012	0.068 (0.029)	0.020
HFE rs1799945	G	0.15	0.001	49	0.065 (0.01)	1.71×10^{-10}	0.005 (0.002)	0.004	0.077 (0.029)	0.007
ABO rs651007	C	0.80	0.001	49	0.05 (0.009)	1.31×10^{-8}	-0.002 (0.002)	0.205	-0.038 (0.031)	0.217

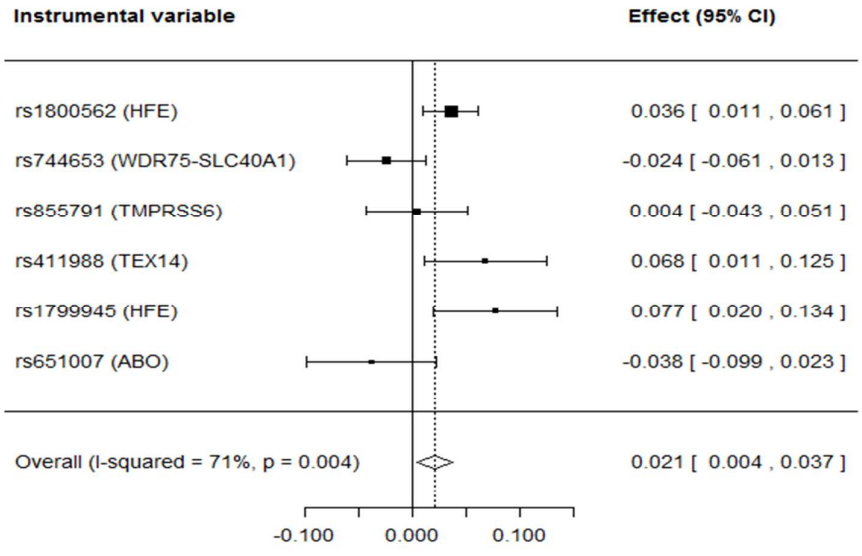
Abbreviations: SNP: single nucleotide polymorphism; EA: effect allele (allele increasing ferritin levels); EA: effect allele frequency; Beta (SE): estimate (standard error) of the association; *p*: p-value; R^2: percentage of the variance of the ferritin explained by the SNP; F: F statistic.

Figure 1: Forest plot for the meta-analysis of MR estimates across variants for: a) *iron-eGFR*; b) *ferritin-eGFR*. The size of the squares is proportional to the precision of the estimate, and the horizontal line indicates its 95%CI; the centre of the diamond represents the pooled MR estimate and the lateral tips indicate its 95% CI.

a) Effect of a SD increase in iron on log(eGFR)



b) Effect of a SD increase in $\log_{10}$(ferritin) on log(eGFR)



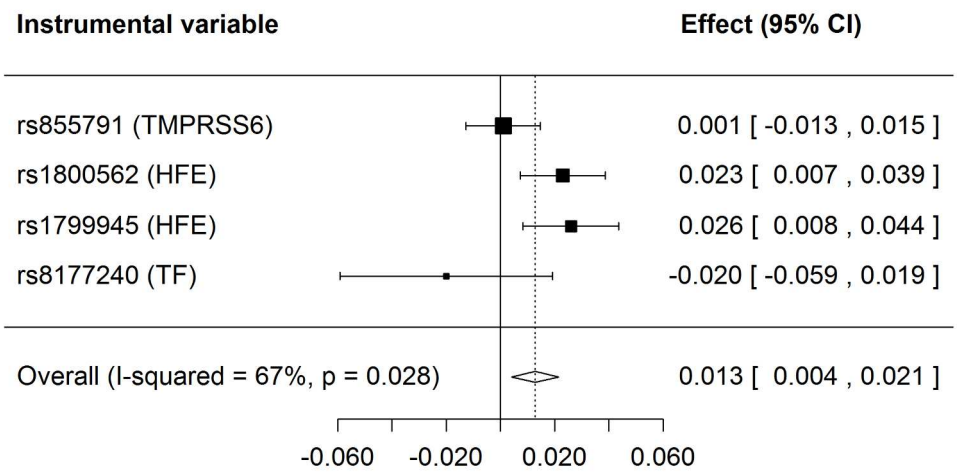


Figure 1:
a) Effect of a SD increase in iron on log(eGFR)
194x110mm (300 x 300 DPI)

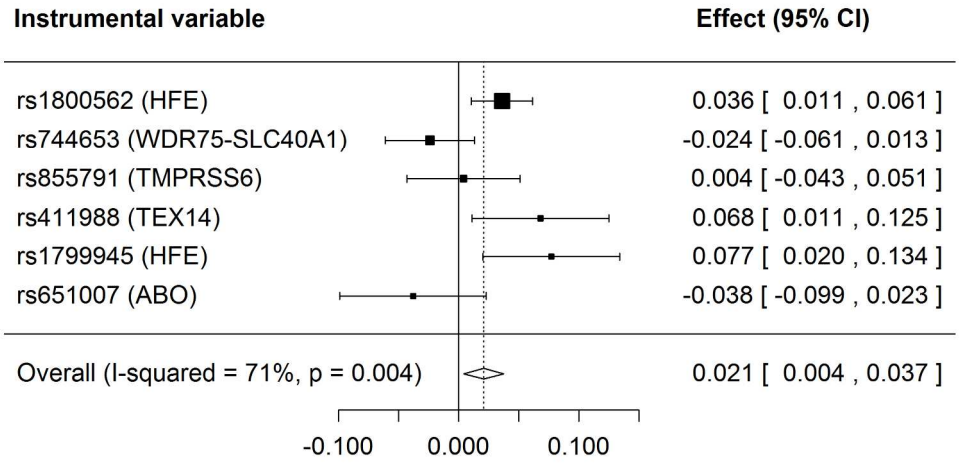


Figure 1:
b) Effect of a SD increase in log10(ferritin) on log(eGFR)
203x110mm (300 x 300 DPI)