medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Causal effects from non-alcoholic fatty liver disease on kidney function: A Mendelian randomization study

Authors

Sehoon Park, MD;[1,2] Soojin Lee, MD;[3,4] Yaerim Kim, MD, PhD;[5] Semin Cho, MD,[3,4] Kwangsoo Kim, PhD;[6] Yong Chul Kim, MD, PhD;[3] Seung Seok Han, MD, PhD;[3,7] Hajeong Lee, MD, PhD;[3] Jung Pyo Lee, MD, PhD;[3,7,8] Kwon Wook Joo, MD, PhD;[3,4,7] Chun Soo Lim, MD, PhD;[3,7,8] Yon Su Kim, MD, PhD;[1,3,4,7] and Dong Ki Kim, MD, PhD[3,4,7]

Author affiliations

[1]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea

[2]Department of Internal Medicine, Armed Forces Capital Hospital, Gyeonggi-do, Korea

[3]Department of Internal Medicine, Seoul National University Hospital, Seoul, Korea

[4]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Korea

[5]Department of Internal Medicine, Keimyung University School of Medicine, Daegu, Korea

[6]Transdisciplinary Department of Medicine & Advanced Technology, Seoul National University Hospital, Seoul, Korea

[7]Kidney Research Institute, Seoul National University, Seoul, Korea

[8]Department of Internal Medicine, Seoul National University Boramae Medical Center, Seoul, Korea

Corresponding author

Dong Ki Kim, MD, PhD

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Professor

Department of Internal Medicine, Seoul National University College of Medicine

101 Daehak-ro, Jongno-gu, Seoul, 03080, Korea

Tel.: 82-2-2072-2303, Fax: 82-2-745-2264

Email: dkkim73@gmail.com

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Abstract

Background & aims: An observational association between nonalcoholic fatty liver disease (NAFLD) and kidney function impairment has been reported. A genetic variant linked to an increased risk of NAFLD, the G allele of rs738409, has been reported to be associated with a reduction in estimated glomerular filtration rate (eGFR).

Methods: In this Mendelian randomization (MR) study, we selected rs738409 as a genetic instrument to predict NAFLD. The genetic variant rs13107325, which is strongly associated with the steatohepatitis marker liver MRI corrected T1 score, was introduced as an alternate genetic instrument for the NAFLD-related phenotype. The eGFR outcome was assessed in individuals of white British ancestry included in the UK Biobank (N = 321,260), and associations adjusted for major metabolic disorder and income grades were investigated. Further, the associations were reassessed in two negative control subgroups (body mass index < 25 kg/m^2 and serum alanine aminotransferase level < 20 IU/mL) with a low probability of developing NAFLD. As a replication analysis, a summary-level MR was performed with the European ancestry CKDGen dataset (N = 567,460).

Results: In the UK Biobank dataset, a genetic predisposition for NAFLD or an increased liver T1 score was significantly associated with a reduced eGFR with adjustment for major metabolic disorders. Although the associations were not significant in the negative control subgroups with a low probability of developing NAFLD, they were significant in the subgroups with a remaining risk of NAFLD, suggesting the absence of a horizontal pleiotropic pathway. The summary-level MR from the CKDGen dataset supported the causal effects of NAFLD on reduced eGFR.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Conclusions: This MR analysis supports the causal reduction in eGFR by NAFLD.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Introduction

Nonalcoholic fatty liver disease (NAFLD) is one of the leading comorbidities worldwide and places a large socioeconomic burden on the population.[1] The prevalence of NAFLD has increased as the population with obesity has grown. Obesity and related metabolic disorders contribute to NAFLD development, and genetic factors also play roles in NAFLD and its severity.

Chronic kidney disease (CKD) is another major comorbidity with a large socioeconomic burden, affecting nearly 7 million individuals in the global population.[2] Previous studies reported that NAFLD may cause kidney function impairment, which is supported by a recent systematic meta-analysis of observational findings.[3] However, as the risks of NAFLD and CKD are both affected by obesity-related metabolic disorders, whether NAFLD has a direct causal effect on kidney function cannot be inferred from conventional observational findings that are inevitably affected by unmeasured confounding effects or reverse causation. Therefore, additional investigation that may evaluate the causal effects of NAFLD on kidney function is warranted. Such evidence would suggest the need to monitor kidney function in NAFLD patients and the possibility that appropriate management of NAFLD may lead to the prevention of kidney function impairment.

Previous studies identified that the PNPLA3 gene single nucleotide polymorphism (SNP), rs738409, is a genetic causal variant for NAFLD.[4] The G allele of rs738409 has been reported to be associated with NAFLD in genome-wide association studies (GWASs) and to cause cirrhotic changes in the liver in NAFLD patients. Interestingly, previous reports identified that rs738409 is also associated with kidney function in several cohorts.[5-8] This may suggest a causal association between the two disorders; however, a further large-scale study is warranted to dissect this association between rs738409 and NAFLD or kidney function to demonstrate the causal direction.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Mendelian randomization (MR) is an analytic tool for estimating the causal effect of an exposure on complex outcomes.[9] As genotype is fixed before birth, genetically predicted exposure is minimally affected by reverse causation or confounding effects. The method has been implemented in the recent medical literature and identified important causal biological pathways between complex disorders. In addition, with the availability of large-scale genetic data for kidney function traits, MR has been utilized to identify causal risk factors for kidney function impairment.[10-12]

In this MR study, we performed MR analysis by selecting rs738409 as a genetic instrument to predict NAFLD and assess its association with kidney function outcome in two independent, population-scale genetic datasets for kidney function traits. We also introduced an alternate genetic instrument, the rs13107325 SNP, which is strongly associated with the liver MRI corrected T1 (cT1) score, which is a marker of steatohepatitis and liver fibrosis,[13] to support the causal linkage between NAFLD and kidney function. We hypothesized that causal effects of NAFLD on kidney function would be present.

Methods

Ethical considerations

This study was approved by the Institutional Review Boards of Seoul National University Hospital (E-2010-007-1160). The requirement for informed consent was waived as we investigated anonymous public databases. The investigation of the UK Biobank was approved by the consortium (No. 53799).

Concept of MR analysis

Unlike direct assessment of the effect of a SNP, MR tests the association between

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

genetically predicted exposure and an outcome, which would be mediated by phenotypical exposure (Figure 1).[9] As not all individuals with a genetic predisposition develop the phenotype of interest, MR requires a large study sample to demonstrate causal estimates. The randomization process occurs before birth, and MR stratifies the population according to their genetic risk for exposure. Regarding the main MR analysis in this study, as the rs738409 G allele has been identified to be the most significant genetic variant for NAFLD, those with the G/G rs738409 genotype would have a higher probability of developing NAFLD than those with the C/C genotype. Thus, if individuals with the rs738409 G allele have worse kidney function than those carrying other alleles, the genetic effect of the G allele on NAFLD would suggest a causal effect of NAFLD on kidney function. As one's genotype precedes the development of clinical confounders or reverse causation, a causal inference of the identified association is possible. To validly demonstrate a causal effect, MR requires attaining three key assumptions: the relevance assumption, independence assumption, and exclusion-restriction assumption. The relevance assumption is that the genetic instrument should be strongly associated with the exposure of interest. The independence assumption and exclusion-restriction assumption indicate the absence of a horizontal pleiotropic pathway, and the assumptions are violated when the genetic effect from an instrument is not through the exposure of interest but through a confounder or another mechanism.

Considerations for MR assumptions

The relevance assumption was considered attained, and instrumented genetic variants had the strongest association with NAFLD or liver function traits in the literature.

The independence assumption was considered attained, as the current study was a single-variant MR study utilizing a SNP with the strongest association with the exposure, and the direct association between the genetic instrument and NAFLD or a related phenotype has

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

been reported in previous studies.[4,13] We selected a single SNP as the genetic instrument because other reported genetic variants associated with NAFLD were mostly associated with obesity-related phenotypes; thus, including them would violate this assumption. Furthermore, when individual-level data were available, we adjusted for major metabolic comorbidities, including hypertension, diabetes mellitus, obesity, dyslipidemia, and income grades, as clinical covariates to demonstrate the independent effect from NAFLD or liver steatohepatitis.

In the current MR analysis, we paid particular attention to the exclusion-restriction assumption, as some reports, despite the inclusion of only a few hundred individuals, suggest that rs738409 may directly affect kidney function.[5,7,8] The previously reported association between rs738409 and kidney function may be explained by three possibilities: 1) rs738409 increases the risk of NAFLD indirectly because NAFLD causes kidney function impairment, affecting kidney function, 2) the effect of NAFLD on kidney function is dependent on rs738409, but the rs738409 effect is accentuated or mediated through NAFLD, or 3) rs738409 affects kidney function independent of its effect on NAFLD. The first and second scenarios would be the attainment of the exclusion-restriction assumption, as the effect is mediated by NAFLD, but the third scenario would be a violation of the assumption, discouraging causal inference from MR.

Currently, there is no direct statistical test assessing the exclusion-restriction assumption. As suggested in the literature, to test whether there is an independent effect of rs738409 on kidney function, we implemented the “negative control approach”, which has been suggested to be useful for testing the absence of a horizontal pleiotropic pathway.[9,14] The negative control approach re-evaluates the association between genetic instruments and outcome in those with a lower probability of developing the exposure phenotype. For instance, if an estimate of the causal effect of genetically predicted alcohol drinking behavior on a health outcome is significant in a female population, this would be less likely to indicate a true

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

causal effect than the presence of a horizontal pleiotropic pathway. To perform this analysis, we constructed two negative control subgroups with a low probability of having the NAFLD phenotype with the available individual-level data. First, we divided subgroups according to whether the participants had a normal range body mass index (< 25 kg/m^2), which is not in a range or overweight or obese, and would have a small probability of having NAFLD. Second, we divided subgroups according to an alanine aminotransferase level < 20 IU/mL. Although the absolute diagnostic criterion with alanine aminotransferase level has not been determined, it would be reasonable that individuals with a low level would less like to have NAFLD. The association between rs738409 genotype and eGFR was reassessed in the negative control group, with minimal risk of NAFLD, and their counterparts, with remaining possibility of NAFLD. If a significant association is identified only in the population with a higher probability of NAFLD, this would support that phenotypical NAFLD is the driving factor of the genetic effects of rs738409 on kidney function. Conversely, if the association is similarly identified or more prominent in the negative controls, this would discourage the interpretation that phenotypical NAFLD is the mediating factor of the genetic effects of rs738409 on kidney function.

Causal linkage between NAFLD and kidney function independent of rs738409

If the MR analysis demonstrates causal linkage between NAFLD and kidney function through MR, the remaining question would be whether the causal effect is dependent on the effect of a single SNP, rs738409. If the causal effect is dependent on rs738409, NAFLD patients with the G allele in the rs738409 genotype would be at risk only of kidney function impairment. Otherwise, if a causal linkage not directly associated with rs738409 is identified, the association between rs738409 and reduced kidney function would be through general effects of NAFLD, which would be a more general finding similar to typical MR findings.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Therefore, we aimed to introduce a genetic instrument predicting NAFLD-related traits independent of the effect of a PNPLA3 variant. As NAFLD is hardly genetically explained independent of rs738409,[4] we considered a proxy-phenotype, liver MRI-derived cT1 value.[13] The liver MRI cT1 value is a noninvasively measured marker for the severity of steatohepatitis and fibrosis. A previous genome-wide association study including 14,440 individuals of European ancestry identified that SNP rs13107325 on the SLC38A8 gene showed the strongest association strength with liver MRI cT1 levels. In addition, as the SLC38A8 gene has been reported to contribute to metal ion transport, encoding a Zrt/lrt-like protein, alteration of the SLC38A8 gene has been reported to cause hepatic injury.[14] Therefore, the rs13107325 SNP was considered a valid instrument for steatohepatitis or liver fibrosis. As NAFLD causes steatohepatitis, a causal effect of steatohepatitis demonstrated by the SNP would support the causal linkage between NAFLD or steatohepatitis and kidney function.

UK Biobank data source

The UK Biobank is a prospective cohort of > 500,000 individuals aged 50-65 years in the United Kingdom.[15,16] The database includes invaluable resources of various phenotypical data, and with deeply genotyped information, the UK Biobank data have been implemented in many genetic studies in the current literature. Regarding this study, the database has a particular strength, as other phenotypical data are available, and direct adjustment for kidney function-related metabolic comorbidities was possible. Moreover, the individual-level data enabled us to perform a negative control analysis, which was crucial for supporting the absence of a horizontal pleiotropic pathway.

MR analysis with the individual-level UK Biobank data

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

In the analysis, we included 337,138 white British UK Biobank participants passing the sample quality control.[11] Among them, 274,353 individuals had available information for baseline eGFR values by the CKD-EPI equation based on serum creatinine values and adjusted clinical covariates. Multivariable linear regression analysis was performed for the eGFR values with the exposure of the minor alleles of rs738409 or rs13107325, with the exposure effect size for NAFLD or cT1 value introduced from previous GWAS studies.[4,13] We constructed a regression model including age, sex, hypertension, diabetes mellitus, obesity, dyslipidemia, and income grades as clinical covariates, identified by self-reports and relevant medication usage history. The analysis in the negative control group and the counterparts was also performed. A two-sided P value < 0.05 was considered indicative of a statistically significant finding.[11]

MR analysis with the CKDGen data

To replicate the findings on a large scale, we demonstrated the causal estimates with the largest GWAS meta-analysis for eGFR performed by the CKDGen consortium.[17] The outcome summary statistics in the database, which are available in the public domain (URL: https://ckdgen.imbi.uni-freiburg.de/), have been utilized for various MR analyses in the current literature regarding kidney function.[11,12] The CKDGen GWAS summary result is independent from the UK Biobank data; thus, a similar finding in the data would be an independent replication of the main analysis.

To minimize the effects of different ethnic distributions, we limited the analysis to individuals of European ancestry (N = 567,460). The summary statistics provided in the data were for log-transformed eGFR traits, which were rescaled to a percent change unit. The effect from addition of the minor alleles of rs738409 or rs13107325 on eGFR was tested by the Wald ratio method, as the current MR analysis instrumented a single SNP, with the

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

exposure effect size for NAFLD or the cT1 value introduced from a previous GWAS. The analysis was performed with the TwoSampleMR package in R (version 3.6.2).[18] The negative control approach analysis was not performed with the CKDGen data, as individual-level data to define subgroups were unavailable.

Results

Characteristics of the data sources

The included UK Biobank participants had a median age of 58 [interquartile range 50, 63] years old, and 48% of them were male. The prevalence of obesity, hypertension, diabetes mellitus, and dyslipidemia was 24%, 20%, 5%, and 17%, respectively. While 5% of the individuals had the highest income grade (> 100000 £ per month), 21%, 26%, 26%, and 22% of individuals reported being in the 52000 to 1000000 £, 31000 to 51999 £, 18000 to 30999 £, and < 18000 £ income categories, respectively.

Individuals of European ancestry in the CKDGen dataset had a median age of 54 years old, and 50% of them were male, with a median eGFR of 91.4 mL/min/1.73 m^2 and a prevalence of CKD of 9% (Supplemental Table 2).

MR results in the UK Biobank data

The G allele of rs738409 was significantly associated with reduced eGFR values in the UK Biobank dataset (Table 1). When the analysis was restricted to those with normal weight (body mass index < 25 kg/m^2) and those with a decreased probability of NAFLD development, the association was attenuated and became nonsignificant. However, when the analysis was performed including those with body mass index ≥ 25 kg/m^2, the G allele of rs738409 was significantly associated with lower eGFR levels. In another negative control

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

group including those with an alanine aminotransferase level < 20 IU/mL, the G allele of rs738409 was nonsignificantly associated with eGFR. On the other hand, in those with alanine aminotransferase ≥ 20 IU/mL, who would have an increased probability of NAFLD, the G allele of rs738409 was significantly associated with reduced eGFR values as the main findings.

Similarly, the T allele of rs13107325, an increased genetically predicted liver cT1 score, was significantly associated with a reduced eGFR. The main findings were that the association was attenuated in participants with a body mass index < 25 kg/m^2 or an alanine aminotransferase level < 20 IU/mL. However, the association was significant in their counterparts with an increased probability of NAFLD development and increased effect sizes.

MR results in the CKDGen data

In the summary-level MR from the independent CKDGen dataset, NAFLD that was genetically predicted by rs738409 was significantly associated with reduced eGFR values (Table 2). Additionally, significant causal estimates were also identified by rs13107325, which genetically predicted an increased liver cT1 value associated with a reduced eGFR value.

Discussion

In this study, MR analysis demonstrated a causal linkage between NAFLD or steatohepatitis and kidney function. When the studied population was stratified according to the presence of the G allele of rs738409, those with an increased genetic predisposition for NAFLD, which would lead to an increased prevalence of NAFLD in these individuals, showed worse kidney function. However, when the analysis was limited to those with a reduced possibility of

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

developing NAFLD, the linkage was not identified, indicating that the genetic effect is actually through NAFLD. The causal linkage between steatohepatitis, which can be caused by NAFLD, and kidney function was also identified, suggesting that this causal linkage is due to the general effects of NAFLD and is not completely dependent only on rs738409.

There have been a number of reports suggesting that the presence of NAFLD is associated with an increased risk of CKD.[3] However, as NAFLD itself primarily develops in individuals with metabolic disorders, which also cause kidney function impairment, the possibility that the findings are affected by reverse causation or residual confounding effects remains because of the inevitable limitation of conventional observational studies. In this study, we performed an MR analysis that is known to be minimally affected by such issues and demonstrated the causal linkage between NAFLD and reduced eGFR. This is a large-scale finding that confirms a previously suggested association between the PNPLA3 variant rs738409 and reduced eGFR in two independent population-scale genetic datasets. In addition, with our efforts to attain the MR assumption, we demonstrated that the causal linkage would occur through NAFLD.

The study paid particular attention to the exclusion-restriction assumption of MR, as previous studies reported that rs738409 may directly affect kidney function independent of the presence of NAFLD.[5,7,8] However, as the majority of studies included NAFLD diagnosed by imaging techniques as an adjustment variable, some undiagnosed or subclinical NAFLD cases might have been present. In addition, as such studies included only a few hundred individuals, the possibility of selection bias remained. As we stated, the presence of an independent effect of rs738409 on kidney function cannot explain the finding that the association was absent in those with a reduced risk of developing NAFLD. In addition, considering that the association strength was much weaker with kidney function than it was with NAFLD, it is reasonable to interpret that the association is mediated by NAFLD.[19] In

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

addition, that rs738409 is associated with reduced kidney function in within those with NAFLD may be because the genetic variant is associated with the severity of NAFLD,[4,13] and severe NAFLD, as suggested by our findings with liver MRI cT1 score, would have a larger effect on kidney function. In addition, as the genetically predicted liver MRI cT1 score also showed important estimates of causal effects on a reduced eGFR, the finding suggests that the linkage between rs738409 and kidney function would be from the direct genetic effect of the G allele of rs738409 on an increased risk of NAFLD and the clinical effect of NAFLD on kidney function.

The causal linkage between liver cirrhosis and kidney function impairment has been suggested in past studies.[20] A hepatic failure state has been recognized to cause "hepatorenal syndrome", and such severe kidney failure due to advanced liver cirrhosis has been reported to be reversible by recovering liver function through transplantation. The suggested mechanisms include loss of effect volume by splanchnic vasodilation, increased inflammatory responses related or unrelated to bacterial translocation, and cirrhotic cardiomyopathy, which can be caused by cirrhosis. As NAFLD is a causal factor for liver cirrhosis in the advanced state, the causal effect of NAFLD on kidney function could be explained by this mechanism and such effect may be present even in the early stages of hepatic cirrhosis. The remaining question is the threshold severity of NAFLD that may significantly affect kidney function, which requires a future clinical study to be answered.

The clinical implication of this study is that, as NAFLD is a causal factor for reduced eGFR, kidney function should be monitored in NAFLD patients. Further, as CKD is an important risk factor for all-cause mortality and cardiovascular diseases,[2] such a state should be promptly diagnosed, and early intervention to delay the progression of kidney function impairment should be considered in NAFLD patients experiencing kidney function decline. In addition, screening for NAFLD may be considered in those with CKD, as progression of

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

NAFLD may accelerate kidney function decline. Individuals with common risk factors, such as obesity, hypertension, or diabetes, may be the particular targets of such screening strategies. Further, future clinical trials may ask whether management of NAFLD is beneficial for the risk of kidney function impairment development.

There are some limitations in this study. First, the currently identified magnitude of the effect size may differ in the clinical situation because of the nature of MR.[21] In addition, as shared risk factors for metabolic comorbidities are common between NAFLD and CKD, the size of the observational association between NAFLD and CKD would be even larger than that reported in the current findings. Second, this study cannot directly prove that appropriate management of NAFLD would reduce the risk of kidney function impairment. Such clinical utility should be shown in a future clinical trial. Third, the included individuals were limited to individuals of European ancestry because the genetic distribution of SNPs may be different according to ethnic background, which would cause bias in the MR analysis. Thus, generalizability should be broadened in a future study investigating other ethnic populations. Last, the construction of a complete negative control was impossible, as NAFLD also occurs in those with a relatively healthy metabolic state. If we stringently define a subgroup with no risk, this would increase the possibility of collider bias, and only a small number of individuals would be included in this subgroup.[22] As the negative controls defined in this study would have a remaining possibility of NAFLD development or worse cT1 score, the findings might have been only marginal, not completely null, in the negative controls. This was true particularly with cT1 exposure because liver fibrosis can occur from etiologies other than NAFLD.

In conclusion, this MR analysis supports that NAFLD causally reduces eGFR. Clinicians should consider the causal linkage between NAFLD and kidney function impairment.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Authors' contributions

The corresponding author attests that all listed authors meet the authorship criteria and that no others meeting the criteria have been omitted. SP, HL, KSK, KWJ, and DKK contributed to the conception and design of the study. SL, YK, SC, YCK, SL, EKC, SSH, JPL, KWJ, CSL, YSK, and DKK provided statistical advice and interpreted the data. SP and KSK performed the main statistical analysis, assisted by SL and YK. HL, JPL, KWJ, CSL, YSK, and DKK provided advice regarding the data interpretation. YCK, SSH, HL, JPL, KWJ, CSL, and YSK provided material support during the study. SP and DKK had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors participated in drafting the manuscript. All authors reviewed the manuscript and approved the final version to be published.

Acknowledgements

The study was based on the data provided by the UK Biobank consortium (application No. 53799). We thank the investigators of the CKDGen, BIOS, and SSGAC consortia who provided the summary statistics for the genetic instrument and the outcome of this study.

Conflicts of interest

None.

Funding

This work was supported by the Industrial Strategic Technology Development Program - Development of biocore technology (10077474, Development of early diagnosis technology for acute/chronic renal failure) funded by the Ministry of Trade, Industry & Energy (MOTIE,

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Korea) and a grant from SNU R&DB Foundation (800-20190571). The study was performed independently by the authors.

Data availability

The data for this study is available in the public domain.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

References

1. Bellentani S. The epidemiology of non-alcoholic fatty liver disease. Liver Int. 2017;37 Suppl 1:81-84.

2. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2020;395:709-733.

3. Mantovani A, Zaza G, Byrne CD, et al. Nonalcoholic fatty liver disease increases risk of incident chronic kidney disease: A systematic review and meta-analysis. Metabolism. 2018;79:64-76.

4. Anstee QM, Darlay R, Cockell S, et al. Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically characterised cohort. J Hepatol. 2020;73:505-515.

5. Oniki K, Saruwatari J, Izuka T, et al. Influence of the PNPLA3 rs738409 Polymorphism on Non-Alcoholic Fatty Liver Disease and Renal Function among Normal Weight Subjects. PLoS One. 2015;10:e0132640.

6. Musso G, Cassader M, Gambino R. PNPLA3 rs738409 and TM6SF2 rs58542926 gene variants affect renal disease and function in nonalcoholic fatty liver disease. Hepatology. 2015;62:658-659.

7. Mantovani A, Zusi C, Sani E, et al. Association between PNPLA3rs738409 polymorphism decreased kidney function in postmenopausal type 2 diabetic women with or without non-alcoholic fatty liver disease. Diabetes Metab. 2019;45:480-487.

8. Sun DQ, Zheng KI, Xu G, et al. PNPLA3 rs738409 is associated with renal glomerular and tubular injury in NAFLD patients with persistently normal ALT levels. Liver Int. 2020;40:107-119.

9. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

10. Park S, Lee S, Kim Y, et al. Causal effects of education on chronic kidney disease: a Mendelian randomization study. Clinical Kidney Journal. 2020.

11. Park S, Lee S, Kim Y, et al. Short or Long Sleep Duration and CKD: A Mendelian Randomization Study. J Am Soc Nephrol. 2020;31:2937-2947.

12. Kennedy OJ, Pirastu N, Poole R, et al. Coffee Consumption and Kidney Function: A Mendelian Randomization Study. Am J Kidney Dis. 2020;75:753-761.

13. Parisinos CA, Wilman HR, Thomas EL, et al. Genome-wide and Mendelian randomisation studies of liver MRI yield insights into the pathogenesis of steatohepatitis. J Hepatol. 2020;73:241-251.

14. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum Mol Genet. 2018;27:R195-r208.

15. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. Am J Epidemiol. 2017;186:1026-1034.

16. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562:203-209.

17. Wuttke M, Li Y, Li M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. Nat Genet. 2019;51:957-972.

18. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018;7.

19. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLoS Genet. 2017;13:e1007081.

20. Ginès P, Solà E, Angeli P, Wong F, Nadim MK, Kamath PS. Hepatorenal syndrome. Nat Rev Dis Primers. 2018;4:23.

21. Burgess S, Butterworth A, Malarstig A, Thompson SG. Use of Mendelian

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

randomisation to assess potential benefit of clinical intervention. Bmj. 2012;345:e7325.

22. Taylor AE, Davies NM, Ware JJ, VanderWeele T, Smith GD, Munafò MR. Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. Econ Hum Biol. 2014;13:99-106.

Table 1. Association between G allele of rs738409 or T allele of rs13107325 in the individual-level UK Biobank data.

Study group	Genetically predicted exposure	Genetic instrument	Adjusted beta (standard error)	P
Total white British ancestry individuals (N = 274,353)	NAFLD	rs738409 (G)	-0.130 (0.063)	0.038
	Liver cT1 score	rs13107325 (T)	-0.326 (0.102)	0.001
Subgroup analysis				
body mass index < 25 kg/m^2 (N = 91,263)	NAFLD	rs738409 (G)	-0.019 (0.102)	0.860
	Liver cT1 score	rs13107325 (T)	-0.167 (0.183)	0.364
body mass index ≥ 25 kg/m^2 (N = 183,090)	NAFLD	rs738409 (G)	-0.204 (0.079)	0.010
	Liver cT1 score	rs13107325 (T)	-0.306 (0.135)	0.023
ALT < 20 IU/mL (N = 133,811)	NAFLD	rs738409 (G)	-0.105 (0.093)	0.250
	Liver cT1 score	rs13107325 (T)	-0.289 (0.159)	0.080
ALT ≥ 20 IU/mL (N = 140,362)	NAFLD	rs738409 (G)	-0.204 (0.085)	0.016
	Liver cT1 score	rs13107325 (T)	-0.368 (0.150)	0.014

NAFLD = non-alcoholic fatty liver disease, cT1 = corrected T1

The exposure unit was scaled to the effects of genetic instruments for predicted exposure which were introduced from previous GWAS studies. The unit of exposure for NAFLD was one exponential increase in odds and for liver cT1 score was for one standard deviation increase and the unit of outcome was continuous eGFR value (mL/min/1.73 m^2)

The causal estimates were calculated by linear regression of the allele increase score reflecting genetically predicted exposures, with adjustment for age, sex, hypertension, diabetes mellitus, obesity, dyslipidemia, and income grades.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252263; this version posted February 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

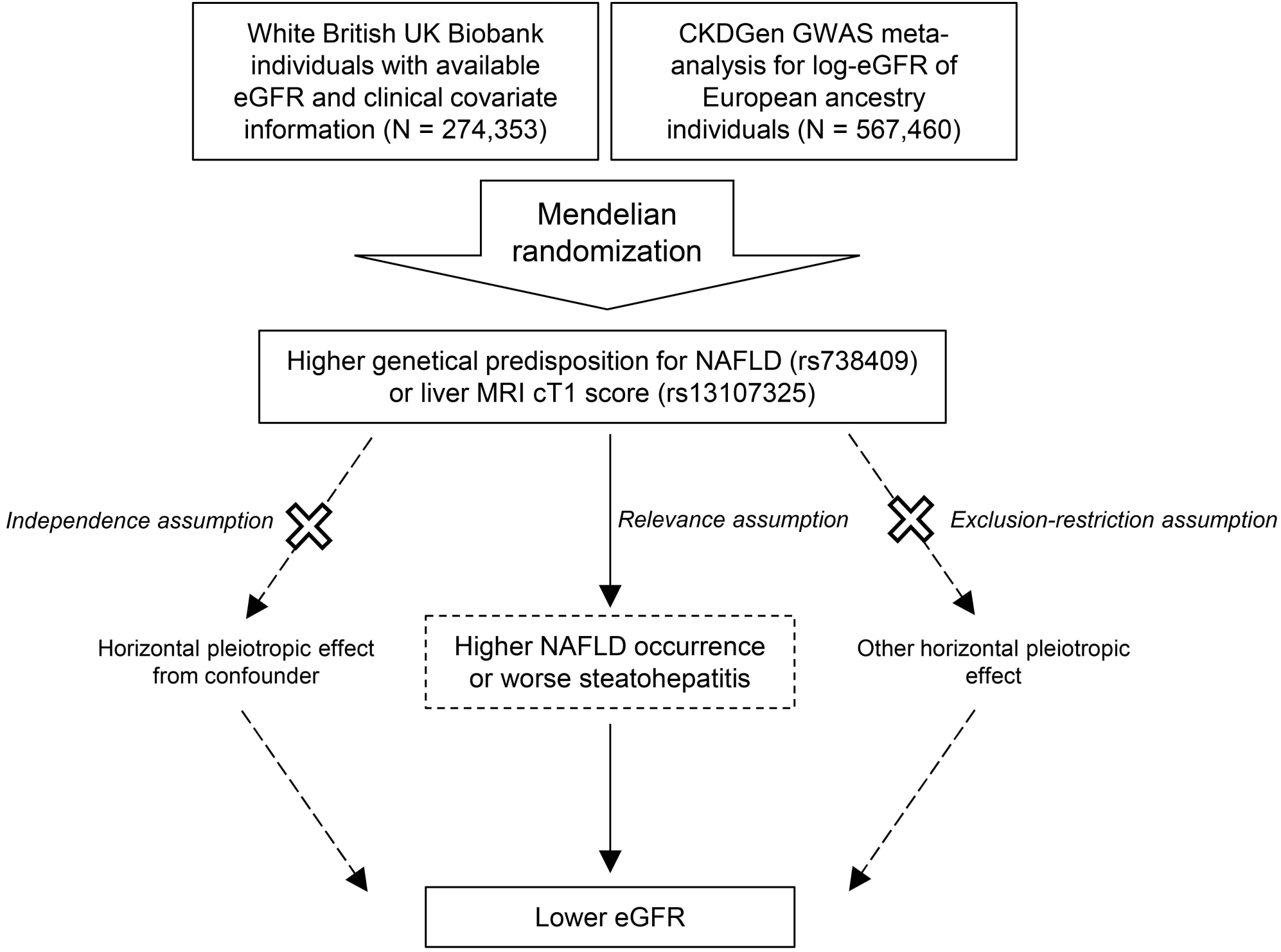
Table 2. Summary-level Mendelian randomization analysis with the European ancestry CKDGen data.

Predicted exposure[a]	Genetic instrument	Beta (standard error) [a]	P
NAFLD	rs738409 (G)	-0.178 % (0.071)	0.012
Liver cT1 score	rs13107325 (T)	-0.515 % (0.143)	< 0.001

NAFLD = non-alcoholic fatty liver disease, cT1 = corrected T1

The exposure unit was scaled to the effects of genetic instruments for predicted exposure which were introduced from previous GWAS studies. The unit of exposure for NAFLD was one exponential increase in odds and for liver cT1 score was for one standard deviation increase and the unit of outcome was % change in eGFR.

[a] The causal estimates were derived from the Wald ratio method.


White British UK Biobank individuals with available eGFR and clinical covariate information (N = 274,353)
CKDGen GWAS meta-analysis for log-eGFR of European ancestry individuals (N = 567,460)
Mendelian randomization
Higher genetical predisposition for NAFLD (rs738409) or liver MRI cT1 score (rs13107325)
Independence assumption
Relevance assumption
Exclusion-restriction assumption
Horizontal pleiotropic effect from confounder
Higher NAFLD occurrence or worse steatohepatitis
Other horizontal pleiotropic effect
Lower eGFR