Supplementary Methods

Details regarding the outcome dataset

The previous CKDGen data had some overlap with the GLGC (Global Lipids Genetics Consortium) study, which provided the genetic instruments, with the possible overlap estimated to be ~30,000 individuals at maximum.

For additional analysis, we used the previous CKDGen phase 4 genome-wide association study (GWAS) meta-analysis summary statistics for log-transformed creatinine-based estimated glomerular filtration rate (eGFR) values, including 567,460 European ancestry data [1,2]. The strength of using the data excluding the UK Biobank data is that this approach can avoid the healthy-volunteer bias of the UK Biobank data [3], as the participants of the UK Biobank had relatively preserved kidney function with low chronic kindney disease prevalence (<5%) compared to the general population. In turn, the issue related to type 1 error emerges because the proportion of sample overlap increases; however, the overall possible number is limited to <10% of the samples included in the outcome data [4].

Next, we used the GWAS summary statistics for cystatin C-based log-transformed eGFR values of the UK Biobank (European ancestry, n = 436,581) [5], as cystatin C-based eGFR values are less affected by diet or muscle mass than creatinine-based values. In addition, using this dataset avoids the sample overlap issue, enabling two-sample Mendelian randomization (MR) analysis with independent datasets.

Last, we also utilized the individual level UK Biobank data of individuals of white British ancestry. The dataset was utilized for nonlinear MR analysis, which requires individual level data with measurements for both exposure and outcome phenotypes [6,7]. A total of 320,598 individuals of white British ancestry who passed genetic quality control with available creatinine-based eGFR values and low-density lipoprotein cholesterol levels were included in the individual level outcome data.

Details regarding the statistical MR analysis methods

We applied the multiplicative random-effects inverse variance-weighted method as the main MR analysis. The method allows balanced pleiotropic effects of the utilized variants and is suggested as the main MR method in the current guidelines [8,9]. Next, the weighted median method, which is one of the representative pleiotropy-robust MR analysis methods, was used [10]. The method yields valid causal estimates even though up to 50% of the instruments are invalid, waiving independence and exclusion-restriction assumptions for half of the instruments at most. MR-Egger regression with bootstrapped standard error was performed with calculation of the MR-Egger intercept p-value [11]. MR-Egger regression provides pleiotropy-robust causal estimates under the attainment of the instrument strength independent of direct effect assumption; however, the weakness of this method is weak statistical power, particularly when the number of instrumented single nucleotide polymorphisms is small. The MR-Egger intercept is commonly used to assess the presence of directional pleiotropy, and if the intercept is significantly different from zero ($p < 0.05$), a possibility of directional pleiotropy was considered to be present.

References

1. 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.
2. Park S, Lee S, Kim Y, et al. Atrial fibrillation and kidney function: a bidirectional Mendelian randomization study. Eur Heart J 2021;42:2816-2823.
3. 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.
4. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol 2016;40:597-608.
5. Stanzick KJ, Li Y, Schlosser P, et al. Discovery and prioritization of variants and genes for kidney function in >1.2 million individuals. Nat Commun 2021;12:4350.
6. Park S, Lee S, Kim Y, et al. Kidney function and obstructive lung disease: a bidirectional Mendelian randomisation study. Eur Respir J 2021;58:2100848.
7. Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental

variables with application to Mendelian randomization. Genet Epidemiol 2017;41:341-352.
8. Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian randomization investigations. Wellcome Open Res 2020;4:186.
9. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. BMJ 2021;375:n2233.
10. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol 2016;40:304-314.
11. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512-525.