

Causal Association between Rheumatoid Arthritis with the Increased Risk of Type 2 Diabetes: A Mendelian Randomization Analysis

Young Ho Lee, Gwan Gyu Song

Department of Rheumatology, Korea University College of Medicine, Seoul, Korea

Objective. This study aimed to examine whether rheumatoid arthritis (RA) is causally associated with type 2 diabetes (T2D). **Methods.** We performed a two-sample Mendelian randomization (MR) analysis using the inverse-variance weighted (IVW), weighted median, and MR-Egger regression methods. We used the publicly available summary statistics datasets from a genome-wide association studies (GWAS) meta-analysis of 5,539 autoantibody-positive individuals with RA and 20,169 controls of European descent, and a GWAS dataset of 10,247 individuals with T2D and 53,924 controls, overwhelmingly of European descent as outcomes. **Results.** We selected 10 single-nucleotide polymorphisms from GWAS data on RA as instrumental variables to improve the inference. The IVW method supported a causal association between RA and T2D ($\beta = 0.044$, standard error [SE] = 0.022, $p = 0.047$). The MR-Egger analysis showed a causal association between RA and T2D ($\beta = 0.093$, SE = 0.033, $p = 0.023$). In addition, the weighted median approach supported a causal association between RA and T2D ($\beta = 0.056$, SE = 0.025, $p = 0.028$). The association between RA and T2D was consistently observed using IVW, MR Egger, and weighted median methods. Cochran's Q test indicated no evidence of heterogeneity between instrumental variable estimates based on individual variants and MR-Egger regression revealed that directional pleiotropy was unlikely to have biased the results (intercept = -0.030 ; $p = 0.101$). **Conclusion.** MR analysis supports that RA may be causally associated with an increased risk of T2D. (*J Rheum Dis* 2019;26:131-136)

Key Words. Rheumatoid arthritis, Type 2 diabetes, Mendelian randomization

INTRODUCTION

Rheumatoid arthritis (RA) is a systemic autoimmune disease characterized by inflammatory changes in the synovial membranes and articular structures, leading to damaged bone structure, causing disability and a decreased quality of life [1,2]. Type 2 diabetes (T2D) is characterized by pancreatic β cell dysfunction and insulin resistance [3]. RA is associated with increased cardiovascular morbidity and mortality [4], and T2D is one of the most important risk factors for cardiovascular disease [5].

Inflammation in RA is characterized by increased levels

of mediators and cytokines (e.g., tumor necrosis factor- α [TNF- α] and interleukin-6 [IL-6]), which is associated with insulin resistance [6]. Several studies have focused on the risk of T2D in RA. A relationship between RA and T2D has been previously reported, although the evidence is inconsistent [7-9]. In a Canadian cohort of 48,718 patients with RA, the risk of T2D compared to a population of 442,033 healthy individuals was increased (hazard ratio 1.5, 95% confidence interval [CI] 1.4~1.5) after adjustment on age, gender, and treatment with disease modifying antirheumatic drugs (DMARDs) or glucocorticoids [7]. In a Taiwan study from 600,695 adults including 4193 with RA, the relative risk (RR) of T2D in the

Received : December 10, 2018, Revised : January 4, 2019, Accepted : January 23, 2019

Corresponding to : Young Ho Lee  <http://orcid.org/0000-0003-4213-1909>

Department of Rheumatology, Korea University Anam Hospital, Korea University College of Medicine, 73 Incheon-ro, Seongbuk-gu, Seoul 02841, Korea. E-mail : lyhcggh@korea.ac.kr

Copyright © 2019 by The Korean College of Rheumatology. All rights reserved.

This is an Open Access article, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

group with RA was 1.68 (95% CI 1.53~1.84) in males and 1.46 (95% CI 1.39~1.54) in females [8]. RA and/or DMARDs used for the treatment of RA may affect the risk of diabetes mellitus by directly altering glucose metabolism; thus, they have been considered risk factors for T2D [10]. Meta-analysis of observational studies have shown that RA is associated with an increased risk of T2D (RR 1.24, 95% CI 1.14~1.35) [11]. However, observational studies are prone to biases, such as reverse causation and residual confounding, thereby precluding a clear understanding of the effect of RA on T2D.

Mendelian randomization (MR) is a technique that uses genetic variants as instrumental variables (IVs) to assess whether an observational association between a risk factor or exposure and an outcome is consistent with a causal effect [12]. A two-sample MR estimates causal effects, in which the exposure and outcome data are measured in different samples [13]. It is very useful in situations where measuring the exposure and outcome in the same set of individuals is difficult [13]. This study aimed to examine whether RA is causally associated with T2D, using a two-sample MR analysis.

MATERIALS AND METHODS

Data sources and selection of genetic variants

We searched the MR Base database (<http://www.mrbase.org/>), which houses a large collection of summary statistic data from hundreds of genome-wide association studies (GWASs). We used a GWAS meta-analysis of 5,539 autoantibody-positive individuals with RA and 20,169 controls of European descent [14] as exposures. A two-sample MR study of genetic variants associated with RA was used as the IV to improve the inference based on a linkage disequilibrium (LD) R^2 of 0.001, clumping distance of 10,000 kb, and p-value threshold of 5.00E-08 (genome-wide significance). We obtained summary statistics (beta coefficients and standard errors) for 10 single-nucleotide polymorphisms (SNPs) associated with RA as the IVs from RA GWAS [14]. We used a T2D GWAS of 10,247 individuals and 53,924 controls, overwhelmingly of European descent [15], as the outcome.

Statistical analysis for Mendelian randomization

MR analysis requires genetic variants to be related to, but not potential confounders of, an exposure [16]. First, we assessed the independent association of 10 SNPs with RA. The 102 SNPs (101 loci) reported to be associated

with RA from the largest published RA GWAS meta-analysis were from subjects of European and Asian ancestries [17]. However, a T2D GWAS dataset was of European descent as outcomes. In order to avoid confounding by ethnicity, we used other European RA GWAS dataset [14]. In addition, genes (polymorphisms) selected for MR should be independent and be available in both the RA and T2D GWAS datasets (exposure and outcome). Therefore, only the 10 genes met the criteria for the MR analyses [14]. Second, we examined the association between each SNP and the risk of T2D. Third, we combined these findings to estimate the uncompounded causal association between RA and the risk of T2D using MR analysis. We performed a two-sample MR, which is a method used to estimate the causal effect of an exposure (RA) on outcomes (T2D) using summary statistics from different GWAS datasets. To assess the causal relationships between RA and the risk of T2D, we used summary data from GWAS on RA and T2D along with their respective 10 SNPs, which were used as IVs.

The IVW method uses a meta-analysis approach to combine the Wald ratio estimates of the causal effect obtained from different SNPs, and provides a consistent estimate of the causal effect of the exposure on the outcome when each of the genetic variants satisfies the assumptions of an IV [18]. Although the inclusion of multiple variants in an MR analysis results in increased statistical power, it has the potential to include pleiotropic genetic variants that are invalid IVs. To explore and adjust for pleiotropy (association of genetic variants with more than one variable), the weighted median and MR-Egger regression methods were performed. MR-Egger regression analysis accounts for the presence of unbalanced pleiotropy by introducing a parameter for this bias, thereby incorporating summary data estimates of the causal effect from multiple individual variants, which is robust to invalid instruments [19]. MR-Egger is a weighted linear regression of the gene-outcome coefficients on the gene-exposure coefficients [19]. The slope of this regression represents the estimate of the causal effect, and the intercept can be interpreted as an estimate of the average horizontal pleiotropic effect across the genetic variants [20]. The weighted median estimator provides a consistent estimate of the causal effect even when up to 50% of the information contributing to the analysis comes from genetic variants that are invalid IVs [21]. Compared to MR-Egger analysis, the weighted median estimator retains greater precision of the estimates [21]. Tests were considered statistically sig-

Table 1. Instrumental single-nucleotide polymorphisms associated with RA and T2D GWASs

Instrumental SNP	Gene	Chromosome loci	Effect allele	Exposure (RA)			Outcome (T2D)		
				Beta	SE	p-value	Beta	SE	p-value
rs10040327	IL6ST	5	A	-0.288	0.044	6.65E-11	-0.037	0.039	0.344
rs3087243	CTLA4	2	A	-0.139	0.023	2.24E-08	0.014	0.023	0.534
rs396458	CYP21A2	6	G	-0.528	0.035	9.15E-50	0.043	0.033	0.195
rs460568	C6orf120	6	T	0.262	0.031	3.81E-16	0.027	0.032	0.391
rs4810485	CD40	20	T	-0.163	0.030	5.69E-09	0.016	0.026	0.538
rs6679677	DCLRE1B	1	A	0.663	0.038	4.39E-70	-0.007	0.036	0.855
rs6920220	LOC102723649	6	A	0.199	0.029	2.49E-12	-0.008	0.028	0.777
rs9268145	C6orf10	6	G	1.058	0.027	<5.00E-08	0.089	0.028	0.001
rs9277412	HLA-DPB1	6	T	-0.357	0.029	1.80E-39	0.014	0.025	0.564
rs9784876	TAP2	6	A	0.668	0.047	2.40E-46	0.067	0.042	0.110

RA: rheumatoid arthritis, T2D: type 2 diabetes, GWAS: genome-wide association study, SNP: single-nucleotide polymorphism, Beta: beta coefficient, SE: standard error, IL6ST: interleukin 6 signal transducer, CTLA4: cytotoxic T-lymphocyte-associated protein 4, CYP21A2: cytochrome P450 family 21 subfamily A member 2, C6orf120: chromosome 6 open reading frame 120, CD40: cluster of differentiation 40, DCLRE1B: DNA cross-link repair 1B, LOC102723649: uncharacterized LOC102723649, C6orf10: chromosome 6 open reading frame 10, HLA-DPB1: major histocompatibility complex, Class II, DP Beta 1, TAP2: transporter 2, ATP binding cassette subfamily B member.

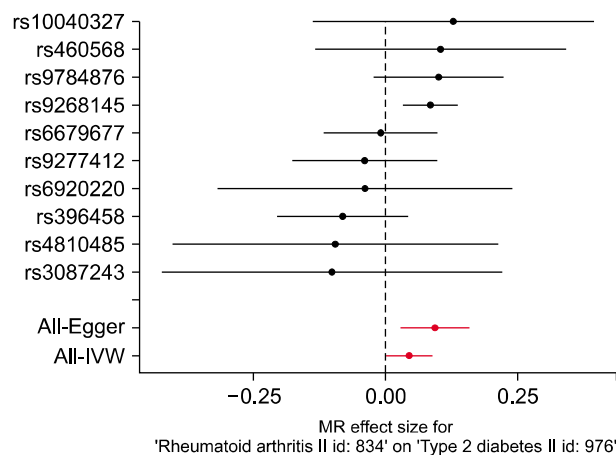


Figure 1. Forest plot of the causal effects of rheumatoid arthritis-associated single-nucleotide polymorphisms on type 2 diabetes. MR: Mendelian randomization, IVW: inverse-variance weighted.

nificant when $p < 0.05$. All MR analyses were performed using the MR-Base platform [22].

Heterogeneity and sensitivity test

We assessed the heterogeneities between the SNPs using Cochran's Q-statistics [23]. The intercept represents the average pleiotropic effect across the genetic variants (the average direct effect of a variant with the outcome). An intercept that differs from zero (the MR-Egger test) indicates directional pleiotropy. We performed MR-Egger regression test to explore pleiotropy.

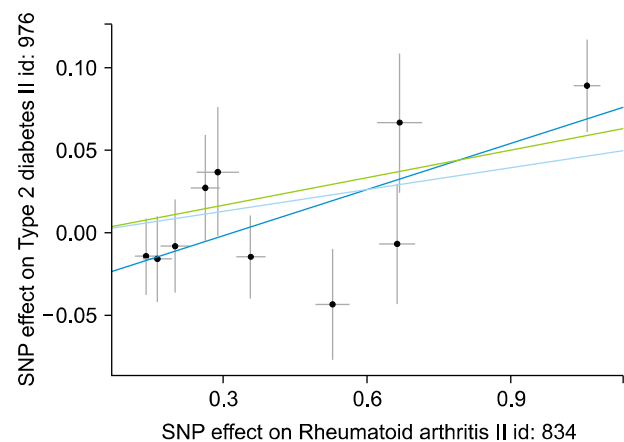


Figure 2. Scatter plots of the genetic associations of rheumatoid arthritis against those of type 2 diabetes. The slopes of each line represent the causal association for each method. Blue line represents the inverse-variance weighted estimate, green line represents the weighted median estimate, and dark blue line represents the MR-Egger estimate. SNP: single-nucleotide polymorphism.

RESULTS

Instrumental variables for Mendelian randomization

We selected 10 independent SNPs from RA GWASs as the IVs; all of them were associated with RA at genome-wide significance (Table 1). Four of the 10 SNPs were positively associated with RA, and the association with rs9268145 was statistically significant (Figure 1).

Mendelian randomization results

IVW method supported a causal association between RA and T2D ($\beta=0.044$, standard error [SE]=0.022, $p=0.047$) (Table 1, Figures 1 and 2). MR-Egger regression revealed that directional pleiotropy was unlikely to have biased the results (intercept = -0.003 ; $p=0.101$). The MR-Egger analysis showed a causal association between RA and T2D ($\beta=0.093$, SE=0.033, $p=0.023$) (Table 1, Figure 2). In addition, the weighted median approach supported a causal association between RA and T2D ($\beta=0.056$, SE=0.025, $p=0.028$) (Table 1, Figure 2). The association between RA and T2D was consistent with the IVW, MR-Egger, and weighted median methods. All these methods suggested a positive causal effect of RA on the risk of T2D. Therefore, the results of MR analysis may support a causal association between RA and T2D.

Heterogeneity and sensitivity test

Cochran's Q test indicated no evidence of heterogeneity between IV estimates based on the individual variants (Table 2). Heterogeneity is the variability in the causal estimates obtained for each SNP (i.e. how consistent is the causal estimate across all SNPs). No evidence of heterogeneity suggests increased reliability of MR estimates. In addition, MR-Egger regression revealed that directional pleiotropy was unlikely to have biased the results (intercept = -0.030 ; $p=0.101$).

DISCUSSION

Most of the well-conducted studies support that RA is associated with the development of T2D [7,8]. The possible link between RA and incident T2D could be explained by the inflammatory responses associated with both conditions. Chronic systemic inflammation plays an important role in the pathogenesis of both RA and T2D [6]. RA is closely linked to the systemic inflammation induced by proinflammatory cytokines such as TNF- α and IL-6, which induce insulin resistance, which leads to T2D by

blocking the function of insulin [24]. Inflammatory pathways are involved in the metabolic abnormalities in T2D. To prove the possible link between RA and incident T2D, genetic variants associated with the inflammatory responses would be selected and used in the MR analysis. However, there is currently no evidence demonstrating that both conditions have shared genetic mechanisms. And the use of DMARDs, such as hydroxychloroquine or tumor necrosis factor inhibitors for lowering the inflammatory response in patients with RA has been shown to reduce the risk of T2D, while glucocorticoid use is associated with an increased risk of T2D by deteriorating glucose tolerance [10].

However, whether RA has a causal relationship with T2D remains unclear, as it was previously reported that the association of T2D with related factors may be a result of bias or confounding factors inherent to observational studies, including glucocorticoid therapy, use of DMARDs that influence the development of T2D, reverse causation, small study numbers and sizes, and selection biases [25,26]. We used three different estimating methods (inverse variance weighting method, weighted median method, and MR-Egger regression) for MR analyses. The MR estimates using IVW, MR-Egger, and weighted median analysis were consistent and supported a causal association between RA and T2D. This MR analysis indicated a causal role of RA in the risk of T2D. Thus, our study corroborates the association found in previous observational studies [7,8].

MR minimizes the possibility of bias inherent to observational studies due to residual confounding or reverse causality [27]. Although the inclusion of multiple variants in MR analyses typically leads to increased statistical power, it also results in the potential inclusion of pleiotropic genetic variants that are invalid IVs [28]. Therefore, approaches used in sensitivity analyses need to be applied to ensure the validity of conclusions drawn from the MR study. To eliminate pleiotropy, we employed a weighted median estimator, which provides valid esti-

Table 2. MR estimates from each method used to determine the causal effect of rheumatoid arthritis on the risk of type 2 diabetes

MR method	Number of SNPs	Beta coefficient	Standard error	Association p-value	Cochran Q-statistic	Heterogeneity p-value
Inverse variance weighted	10	0.044	0.022	0.047	12.14	0.206
MR-Egger	10	0.093	0.033	0.023	8.50	0.387
Weighted median	10	0.056	0.025	0.028	NA	NA

MR: Mendelian randomization, SNP: single-nucleotide polymorphism, NA: not available.

mates even when 50% SNPs are invalid instruments [21]; in addition, we used MR-Egger regression to test for unbalanced pleiotropy and to determine the influence of the causal estimate exposure on the outcome [19]. The MR-Egger approach showed no evidence for unbalanced pleiotropy or a causal association between RA and T2D. Our weighted median estimator results were similar to those obtained using the IVW and MR-Egger approaches, thereby strengthening our confidence in these associations. Our data support previous observational studies that have shown a link between RA and T2D. The current findings may provide an opportunity to determine the underlying mechanisms of the effects of RA on the risk factors of T2D. However, limitations of the two-sample MR must be acknowledged [13]. The disadvantages of using summary data in two sample MR are similar to those of meta-analysis of summary data of randomized controlled trials; the quality of the pooled results is dependent on that of the individual studies. Also the following potential biases limit the evidence for causality from studies; pleiotropy which occurs when a genetic variant influences multiple phenotypic traits or has multiple biological effects, and canalization which refers to the compensatory effects of mechanisms that act to buffer any genetic effect in utero or after birth, and overlap with datasets of different size which is likely to lead to substantial bias in a “two-sample”, and lack of suitable polymorphisms for studying modifiable exposures of interest.

The present study has several limitations. First, genetic variants have a modest effect on RA, as they explain only a small proportion of variance in a particular exposure [29]. This means that a very large number of cases are required to detect a causal relationship for the outcome of interest. Second, the study on T2D was based on participants of European ancestry. As causality may depend on ethnicity and selection bias, further MR studies are required in other populations for a better holistic understanding. Nevertheless, this meta-analysis also has strength. Unlike observational studies on associations between RA and higher risk of T2D, our study was not susceptible to bias from unmeasured confounders or reverse causation by adopting MR approaches and demonstrated that RA may play a causal role in the etiology of T2D. Although RA has been studied as a potential risk factor for T2D, an MR study has been never performed. To the best of our knowledge, this is the first MR study regarding the causal relationship between RA and T2D.

CONCLUSION

In conclusion, the MR analysis suggests that RA may be causally associated with an increased risk of T2D. MR estimates using IVW, weighted median, and MR-Egger analysis were consistent, and the results of MR analysis supported that RA was causally associated with an increased risk of T2D. The current findings may provide an opportunity to determine the underlying mechanisms of the effects of inflammation on the risk of T2D.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Edwards CJ, Cooper C. Early environmental factors and rheumatoid arthritis. *Clin Exp Immunol* 2006;143:1-5.
2. Lee YH, Bae SC, Choi SJ, Ji JD, Song GG. Genome-wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis. *Mol Biol Rep* 2012;39:10627-35.
3. Prentki M, Nolan CJ. Islet beta cell failure in type 2 diabetes. *J Clin Invest* 2006;116:1802-12.
4. Avina-Zubieta JA, Thomas J, Sadatsafavi M, Lehman AJ, Lacaille D. Risk of incident cardiovascular events in patients with rheumatoid arthritis: a meta-analysis of observational studies. *Ann Rheum Dis* 2012;71:1524-9.
5. Shah AD, Langenberg C, Rapsomaniki E, Denaxas S, Pujades-Rodriguez M, Gale CP, et al. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people. *Lancet Diabetes Endocrinol* 2015;3:105-13.
6. Nicolau J, Lequerré T, Bacquet H, Vittecoq O. Rheumatoid arthritis, insulin resistance, and diabetes. *Joint Bone Spine* 2017;84:411-6.
7. Solomon DH, Love TJ, Canning C, Schneeweiss S. Risk of diabetes among patients with rheumatoid arthritis, psoriatic arthritis and psoriasis. *Ann Rheum Dis* 2010;69:2114-7.
8. Su CC, Chen IeC, Young FN, Lian IeB. Risk of diabetes in patients with rheumatoid arthritis: a 12-year retrospective cohort study. *J Rheumatol* 2013;40:1513-8.
9. Simard JF, Mittleman MA. Prevalent rheumatoid arthritis and diabetes among NHANES III participants aged 60 and older. *J Rheumatol* 2007;34:469-73.
10. Ozen G, Pedro S, Holmqvist ME, Avery M, Wolfe F, Michaud K. Risk of diabetes mellitus associated with disease-modifying antirheumatic drugs and statins in rheumatoid arthritis. *Ann Rheum Dis* 2017;76:848-54.
11. Jiang P, Li H, Li X. Diabetes mellitus risk factors in rheumatoid arthritis: a systematic review and meta-analysis. *Clin Exp Rheumatol* 2015;33:115-21.
12. Burgess S, Daniel RM, Butterworth AS, Thompson SG;

- EPIC-InterAct Consortium. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. *Int J Epidemiol* 2015;44:484-95.
13. Lawlor DA. Commentary: two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol* 2016;45:908-15.
 14. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet* 2010;42:508-14.
 15. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segrè AV, Steinthorsdottir V, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;44:981-90.
 16. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;37:658-65.
 17. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;506:376-81.
 18. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol* 2013;178:1177-84.
 19. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512-25.
 20. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* 2017;32:377-89.
 21. 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;40:304-14.
 22. Hemani G, Zheng J, Wade KH, Laurin C, Elsworth B, Burgess S, et al. MR-Base: a platform for systematic causal inference across the phenome using billions of genetic associations. *bioRxiv* 2016. DOI: 10.1101/078972.
 23. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ* 1997;315:1533-7.
 24. Popa C, Netea MG, van Riel PL, van der Meer JW, Stalenhoef AF. The role of TNF-alpha in chronic inflammatory conditions, intermediary metabolism, and cardiovascular risk. *J Lipid Res* 2007;48:751-62.
 25. Song GG, Bae SC, Lee YH. Association between vitamin D intake and the risk of rheumatoid arthritis: a meta-analysis. *Clin Rheumatol* 2012;31:1733-9.
 26. Viswanathan M, Berkman ND, Dryden DM, Hartling L. Assessing risk of bias and confounding in observational studies of interventions or exposures: further development of the RTI item bank. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013 Aug. Report No.: 13-EHC106-EF 2013.
 27. Smith GD, Ebrahim S. Mendelian randomization: genetic variants as instruments for strengthening causal inference in observational studies. In: Weinstein M, Vaupel JW, Wachter KW, eds. *Biosocial surveys*. Washington (DC), National Academies Press (US), 2008.
 28. Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol* 2004;33:30-42.
 29. Swerdlow DI, Kuchenbaecker KB, Shah S, Sofat R, Holmes MV, White J, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. *Int J Epidemiol* 2016;45:1600-16.