

Causal effects from tobacco smoking on obesity-related traits: a Mendelian randomization study

Authors

Sehoon Park, MD;¹ Seong Geun Kim, MD;² Soojin Lee, MD;^{3,4} Yaerim Kim, MD, PhD;⁵ Semin Cho, MD;⁶ Kwangsoo Kim, PhD;⁷ Yong Chul Kim, MD, PhD;² Seung Seok Han, MD, PhD;^{2,8} Hajeong Lee, MD, PhD;² Jung Pyo Lee, MD, PhD;^{4,8,9} Kwon Wook Joo, MD, PhD;^{2,4,8} Chun Soo Lim, MD, PhD;^{4,8,9} Yon Su Kim, MD, PhD;^{1,2,4,8} and Dong Ki Kim, MD, PhD^{2,4,8}

Author affiliations

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

²Department of Internal Medicine, Seoul National University Hospital, Seoul, Korea

³Department of Internal Medicine, Uijeongbu Eulji University Medical Center, Seoul, Korea

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

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

⁶Department of Internal Medicine, Chung-Ang University Gwangmyeong Hospital, Gyeonggi-do, Korea

⁷Transdisciplinary Department of Medicine & Advanced Technology, Seoul National University Hospital, Seoul, Korea

⁸Kidney Research Institute, Seoul National University, Seoul, Korea

⁹Department of Internal Medicine, Seoul National University Boramae Medical Center, Seoul, Korea

Address for correspondence

Dong Ki Kim, MD, PhD

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

Short title: Smoking and obesity

Word counts

Abstract: 236 words

Main text: 2827 words

Keywords: body mass index, waist, tobacco, cigarette, obesity, Mendelian randomization, public health

Abstract

Background: There is a notion that tobacco smoking would have weight control effect based on the appetite suppressive effect of nicotine. However, a causal effect from being an ever smoker on obesity-related traits in the general population has yet been determined.

Methods: This Mendelian randomization (MR) analysis instrumented 378 genetic variants associated with being an ever smoker which mostly initiated in adolescents or young adulthood, identified from a genome-wide association study (GWAS) meta-analysis of 1.2 million individuals. The outcome data for body mass index, waist circumference, hip circumference, and waist-to-hip ratio was collected in 337,318 white British ancestry UK Biobank participants with 40-69 ages. Replication analysis was performed for GWAS meta-analysis for body mass index including the GERA/GIANT data including 364,487 mostly European samples. Summary-level MR by inverse variance weighted method and pleiotropy-robust MR methods, including median-based and MR-Egger regression, was performed.

Results: Summary-level MR analysis indicated that genetically predicted being an ever smoker is causally linked to higher body mass index [$+0.28$ (0.18, 0.38) kg/m^2], waist circumference [$+0.88$ (0.66, 1.10) cm], hip circumference [$+0.40$ (0.23, 0.57) cm], and waist-to-hip ratio [$+0.006$ (0.005, 0.007)]. The results were consistently supported by pleiotropy-robust MR analysis. In the replication analysis, genetically predicted being an ever smoker was again significantly associated with higher body mass index [$+0.03$ (0.01, 0.05) kg/m^2].

Conclusion: Initiation of tobacco use may consequently lead to worse obesity-related traits of the general population in middle-to-old ages.

Background

Tobacco smoking is an important global health problem which increases risks of cardiovascular disease, malignancy, and death.¹ Tobacco smoking is responsible for more than 8 million deaths a year, and nearly 1.2 million individuals die related to second-hand smoking. Medical society pays large efforts to reduce tobacco use, and some promising results suggests the socioeconomic burden related to smoking might have started to decrease.

There has been a notion that tobacco smoking reduces weight.² The concept of slimness related to tobacco use has been even advertised by tobacco companies and the misconception considering tobacco as weight control measure led young adolescents to initiate smoking.^{3,4} That nicotine suppresses appetite in short-term and weight gain phenomenon after smoking cessation even consolidated the misbelief that smoking would be helpful to reduce weight.^{5,6} On the other hand, recent studies demonstrated that being obese causes smoking initiation, suggesting the causal linkage between smoking and obesity.⁷ Yet, the causal effects from smoking initiation on obesity-related traits in the general population warrants additional study. A population-scale evidence testing whether smoking initiation truly leads to lower weight in middle-to-old age individuals would be important for public policy. Furthermore, if the notion turns out to be false, the information should be widely used for social education to prevent tobacco initiation in the vulnerable population who may use tobacco anticipating the weight control effect.³

Mendelian randomization (MR) is an analytic tool to investigate causal effects implementing genetic instrumental variable.⁸ MR has been widely used in medical literature testing the causal estimates from modifiable exposure on health outcomes. As genetic instrumental variable is fixed before birth, MR can demonstrate causal estimate minimally biased from confounding or reversal causal effects.

In this study, we aimed to test the causal effects from tobacco smoking on obesity-related traits. We performed MR analysis in two-large scale genetic databases. We found that being an ever smoker may consequently leads to higher body mass index and worse obesity-related characteristics in middle-to-old age populations.

Methods

Ethical considerations

The study was approved by the institutional review boards of Seoul National University Hospital (No. E-2203-053-1303) and the UK Biobank consortium (application No. 53799).^{9,10} The study was performed in accordance with the Declaration of Helsinki. The requirement for informed consent was waived by the attending institutional review boards because public anonymous database collected for research purpose was used for this study.

Genetic instruments for tobacco smoking

We used the findings from one of the largest GWAS meta-analysis for over 1.2 million European ancestry individuals including over half a million ever smokers (Figure 1).¹¹ In the meta-analysis, 29 studies were included and the representative studies which provided relatively large number of samples were 23andMe, UK Biobank, deCODE, HUNT and others. Total of 378 lead single nucleotide polymorphisms (SNPs) with conditionally genome-wide significant association with being an ever smoker was used as the genetic instrumental variable in this study (Supplementary Table 1). The mean age of tobacco smoking initiation ranged from 13.7 to 22.9 years old in the dataset, indicating the study analyzed being an ever smoker which mostly initiated in adolescents or in early periods of adulthood as in the general population.

MR assumptions

A valid MR analysis should attain three core assumptions to demonstrate causal effects.⁸ First, the relevance assumption is that the genetic instrumental variable should be strongly associated with the exposure of interest. We collected genetic variants with genome-wide significant association with the tobacco smoking phenotype, explaining ~2% variance for smoking initiation,¹¹ thus, the assumption was considered attained. Second, the independence assumption is that the instrumental variable should not be associated with a confounder. We performed pleiotropy-robust MR analysis which can relax the assumption and statistically tested the possibility of a directional pleiotropy (e.g. MR-Egger regression).¹² Also, we performed sensitivity analysis with the scanned results using the GWAS catalog to screen any potentially confounder-associated SNPs.^{11,13} We performed the sensitivity analysis after excluding the SNPs with identified significant association with other traits. Third, the exclusion-restriction assumption is that the tested causal effect should be through the exposure of interest. Although this assumption is not statistically testable, we performed MR analysis by median-based methods which relaxes the assumption for up to half of the instruments, serving as a sensitivity analysis.¹⁴

Steiger filtering

That obesity causally affects tobacco smoking behavior has been reported previously.⁷ In such situation, reverse directional genetic effect should be carefully considered for an MR analysis. Namely, a genetic variant that has stronger association with an outcome phenotype than with an exposure may reflect the reverse genetic effect from outcome to exposure, thus, should be removed from the genetic instrumental variable. Steiger filtering provides the statistical method to test the direction of the effect from a genetic variant, comparing the

explained variance for exposure and for outcome phenotypes.¹⁵ We used the process with the UK Biobank data and the SNPs reflecting genetic effects with direction from tobacco smoking on obesity-related traits were implemented in this study.

UK Biobank outcome dataset and phenotypes

The UK Biobank data was implemented as the main outcome dataset for this study (Figure 1).^{9,10} The main strength of the data is availability of various anthropometric parameters related to obesity and the large-scale genetic information for robustly quality controlled samples. The limitation of the dataset for this study is that the developmental GWAS meta-analysis for the genetic instrumental variable includes the UK Biobank data, and such sample overlap may cause bias towards observational confounding when instrumental strength is weak.¹⁶

Among the > 500,000 participants, we collected the information from 337,318 white British ancestry samples who passed sample quality control, and those who were outliers regarding heterogeneity, missing rate, with excess kinship, or sex chromosomal aneuploidy were excluded.^{17,18}

We collected body mass index, waist circumference, hip circumference and waist-to-hip ratio which were measured at the baseline visits when the participants had age from 40 to 69. To calculate the association between genetic instruments and outcome phenotype, linear regression analysis was performed for the SNPs by PLINK 2.0 and the genetic effect sizes were adjusted for age, sex, age×sex, age², and 10 genetic principal components.¹⁹

Replication outcome dataset

In consideration for the sample overlap between the GWAS meta-analysis for tobacco smoking and the UK Biobank data, an independent replication outcome data was warranted.

Two-sample MR with lower proportion of sample overlap has strength in that the analysis is less likely to be biased towards observational associations.¹⁶

The previous GWAS meta-analysis for body mass index using the Genetic Epidemiology Research on Adult Health and Aging (GERA) and Genetic Investigation of ANthropometric Traits (GIANT) consortium was used as the replication outcome data, which was available in the GWAS catalog (URL: <https://www.ebi.ac.uk/gwas/studies/GCST006368>) (Figure 1).^{13,20} The meta-analysis included 364,487 samples mostly of European ancestry towards body mass index phenotype (Supplementary Table 2). The GERA data was not included in the previous meta-analysis which provided the genetic instrumental variable, as GERA data was only used for GWAS towards drinking behavior in the meta-analysis.^{11,21} The GIANT data had some sample overlap (e.g. ARIC, EGCUT, HRS, METSIM, QIMR studies) with the GWAS meta-analysis for tobacco smoking,²² yet the number of samples included in the both dataset was ~50,000 at maximum, which is much lower than the UK Biobank data has.

Regarding other obesity-related phenotypes, a summary statistic with large-scale sample size or sufficient number of genotyped SNPs was scarce, thus, we performed the replication analysis towards body mass index, which would be the primary parameter that defines obesity.

Summary-level MR analysis

Among the 378 SNPs instruments, only a SNP (rs3076896) was not available in the UK Biobank data which was not included in the according analysis. However, larger number of SNPs were unavailable in the GERA/GIANT GWAS meta-analysis for body mass index, thus, we used proxy SNPs (linkage disequilibrium $R^2 > 0.8$) to replace the missing. This process allowed us to include all SNPs except two (rs3076896, rs748832) among the variants passing the Steiger filtering process in the UK Biobank for body mass index trait. Next, the summary-

level datasets were harmonized and the palindromic SNPs with intermediate allele frequencies were disregarded because of ambiguous strandness.

Multiplicative random-effect inverse variance weighted method was used to calculate main causal estimates following the current guideline for MR. Additional sensitivity analysis was performed including weighted median and simple median methods, which yields valid causal estimates even up to half of the instruments are invalid. The weighted median method allows invalidity in up to half of the instrumented weights, while the simple median method in up to half of the variants, having distinct strengths as pleiotropy-robust MR sensitivity analysis. MR-Egger regression with bootstrapped standard error was performed which relaxes the independence assumption under attainment of the InSIDE assumption. MR-Egger intercept P value was calculated which statistically tests the possibility of directional pleiotropy, and a significantly non-zero intercept (MR-Egger intercept P value < 0.05) was considered to indicate a possible pleiotropic effect. Bonferroni-adjusted two-sided P value of < 0.05/5 by the inverse variance-weighted method was used to indicate statistical significance. The causal estimates were scaled to the effects from two-fold increase in prevalence of being a smoker by multiplying 0.693 to the raw causal estimates.²³ The summary-level MR analysis was performed by “TwoSampleMR” package in R (version 3.6.0) and the MR-base web platform.²⁴

Results

Characteristics of the outcome datasets

The final UK Biobank outcome data had median age of 58 years old (interquartile range 51, 63) (Table 1). Male sex was identified for 46.3%, median body mass index was 26.7 (interquartile range 24.1; 29.9) kg/m². The median values for waist circumference, hip

circumference, and waist-to-hip ratio were 90 (80, 99) cm, 102 (97, 108) cm, and 0.874 (0.802, 0.937), respectively. In the data, the prevalence of diabetes mellitus was 4.8% and total of 20.9% and 17.5% of the samples were taking hypertension and dyslipidemia medications, respectively.

The GERA GWAS had mean ages of ~60 years old including ~ 60% male sex. Their mean body mass index was 28.0 kg/m² for male and 27.3 kg/m² for female in the European ethnic population. The study included in the GIANT meta-analysis for body mass index had median age ranges from 40s to 70s. Their median body mass index among studies ranged from 21.1 to 35.5 in males and 20.4 to 37.9 in females.

MR results in the UK Biobank

After Steiger filtering and harmonization process, 282, 292, 306, and 299 SNPs were used in the MR analysis towards body mass index, waist circumference, hip circumference and waist-to-hip ratio (Supplementary Table 1). The remaining SNPs had genome-wide significant association with tobacco smoking initiation but were weakly bond to outcome phenotypes, ensuring the direction of the MR analysis is from the genetically predicted tobacco use towards obesity-related traits.

The main causal estimates by multiplicative random-effect inverse variance weighted method provided consistently significant results from higher prevalence of tobacco smoking on higher body mass index [from two-fold increase in prevalence, +0.28 (0.18, 0.38) kg/m²], higher waist circumference [+0.88 (0.66, 1.10) cm], higher hip circumference [+0.40 (0.23, 0.57)], and higher waist-to-hip ratio [+0.006 (+0.005, 0.007)] (Figure 2 and Table 2).

The causal estimates remained significant towards all studied outcomes in the UK Biobank data by the median-based methods with the same directional effect. The results were even supported by MR-Egger regression with bootstrapped standard error with similar effect sizes

towards body mass index [+0.23 (0.05, 0.42) kg/m²], waist circumference [+0.85 (0.41, 1.29) cm], and waist-to-hip ratio [+0.006 (0.004, 0.007)]. The result towards hip circumference was attenuated [+0.26 (-0.06, 0.58) cm], however, MR-Egger intercept P value indicated low probability of directional pleiotropy throughout the analysis (intercept P > 0.05).

The results were similarly identified in our sensitivity analysis after disregarding the SNPs that were identified to have notable association with other traits in the GWAS catalog (Supplementary Table 3), and genetically predicted tobacco smoking was consistently associated with worse obesity-related traits.

MR results in the GERA/GIANT data

When applying the Stieiger-filtered SNPs towards the GERA/GIANT data, 275 SNPs were utilized to genetically predict tobacco smoking initiation (Supplementary Table 2). The MR analysis demonstrated that genetically predicted tobacco smoking was significantly associated with higher body mass index [+0.03 (0.01, 0.05)], supported by the weighted median [+0.03 (0.00, 0.05)] and the simple median [+0.03 (0.01, 0.05)] methods (Figure 2 and Table 3). MR-Egger regression with bootstrapped error also yielded significant causal estimates [+0.04 (0.00, 0.08)], and MR-Egger intercept P value indicated low probability of a directional pleiotropy (intercept P = 0.40). The results were similarly reproduced when the SNPs with potential association with other traits screened in the GWAS catalog (Supplementary Table 4)

Discussion

In this MR analysis, we identified that genetically predicted initiation of tobacco smoking was significantly associated with worse obesity-related traits. Considering the data

characteristics, our study suggests that being an ever smoker, mostly initiated in adolescents or in the early adulthood, would causally lead to higher body mass index, higher waist circumference, and higher waist-to-hip ratio in the later life period. Our study supports that tobacco smoking may consequently lead to worse obesity-related traits and that the misbelief for weight control effect from tobacco smoking may not hold in long-term.

The causal linkage between obesity and tobacco use has been suggested before.⁷ The previous study reported that being obesity itself is a risk factor to initiate tobacco use. Obesity-related social disadvantages may be associated with the initiation of tobacco use in the population, and the findings supported the close association between obesity and tobacco use in a previous observational report.²⁵ On the other hand, the causal effects from being an ever smoker in obesity-related traits in the middle-to-old age population have been yet determined by MR analysis, which is a useful tool to investigate causal effects from modifiable exposure on complex outcomes. A previous study addressed that high tobacco consumption may lower weight by MR analysis,²⁶ however, the message is limited to current smokers and a single-SNP analysis is easy to be affected from horizontal pleiotropic effect and weak instrumental bias.²⁷ Our study has strengths in that 1) we used a strong genetic instrumental variable towards being an ever smoker based from previous large-scale GWAS meta-analysis, 2) we paid particular attention to the reverse directional genetic effect by Setiger filtering, 3) potential confounder effects by pleiotropy-robust sensitivity analyses, 4) and replicated the findings in another independent outcome summary-level data. With the consistent results, our MR results supported that being an ever smoker would be a causal risk factor for obesity.

The notable point is that the smoking initiation mostly occurred in adolescents or in young adulthoods and the obesity-related traits were measured in middle-to-old age individuals. Thus, this study suggests that smoking initiation in early period of life would consequently

lead to increased body weight in long-term. The evidence is important to educate the young population not to start tobacco use, particularly affected from the misbeliefs that tobacco smoking may have weight control effect.^{3,4} Although the short-term appetite suppression effect from nicotine would be present,⁵ tobacco smoking may lead to unhealthy behaviors or social factors contributing to obesity in long-term.^{28,29} In addition, having no history of tobacco smoking would be helpful for physical activity and other healthy lifestyles, supporting the identified causal effects from being an ever smoker on higher body mass index and waist circumference.³⁰ Thus, being non-smoker would be the best way to lower the risk of obesity in later period life and people should not initiate tobacco use in anticipation for long-term weight control effect, which also elevates the risk of diabetes and metabolic syndrome.^{31,32} In addition, considering that obesity is one of the most important risk factors for cardiovascular diseases and mortality in modern medicine, policies should also consider reduction in tobacco initiation rate in the population would finally ameliorate the substantial socioeconomic burden related to obesity.

There are limitations in the study. First, the outcome data had some overlap with the dataset provided the genetic instrumental variable.¹⁶ However, in the replication analysis, the proportion was small and considering the previous simulation study, the potential bias would not substantially influence the findings. Second, performing MR analysis with binary exposure tests the association between genetic liability and outcomes.²³ However, utilizing quantitative traits (e.g. smoking heaviness) as exposure phenotypes was not possible for tobacco smoking as testing the causal estimates only in a subgroup (e.g. smokers) causes collider bias, thus the effect within smokers would require different study design.³³ Third, MR analysis cannot determine the mediating mechanism from being an ever smoker on higher body mass index. There may be various vertical pathways (e.g. healthy lifestyles related to being a non-smoker or social factors in tobacco users) which warrants additional

study. Lastly, the study was performed mostly within European ancestry individuals, thus the generalizability to other ethnic population should be expanded in the future.

In conclusion, genetically predicted being an ever smoker causally leads to worse obesity-related traits in middle-to-old age populations. Young people should be educated that initiation of tobacco use would consequently lead to higher risk of obesity in the later periods of life. Healthcare providers and policy makers should mind the adverse effect from tobacco initiation on the burden of obesity and move forward to prevent the people from starting tobacco use.

Authors' contribution

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, SK, YK, YCK, SSH, JPL, KWJ, CSL, YSK, and DKK provided statistical advice and interpreted the data. SP, KSK performed the main statistical analysis, assisted by SK, 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 consortium who provided the summary statistics for the outcome data of this study.

Conflicts of interest

None.

Funding

This research was supported by a grant of the MD-Phd/Medical Scientist Training Program through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT, Ministry of Science and ICT) (No. 2021R1A2C2094586). The study was performed independently by the authors.

Data availability statement

The data used for this study is available in the public database. The CKDGen data is openly available in the consortium website (URL: <https://ckdgen.imbi.uni-freiburg.de/>) and the UK Biobank data is accessible after acquiring approval from the organization (application No. 53799).

References

1. Ghebreyesus TA. Progress in beating the tobacco epidemic. *Lancet*. 2019;394:548-549.
2. Chiolero A, Faeh D, Paccaud F, Cornuz J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. *Am J Clin Nutr*. 2008;87:801-809.
3. Honjo K, Siegel M. Perceived importance of being thin and smoking initiation among young girls. *Tob Control*. 2003;12:289-295.
4. Amos A, Haglund M. From social taboo to "torch of freedom": the marketing of cigarettes to women. *Tob Control*. 2000;9:3-8.
5. Chen H, Vlahos R, Bozinovski S, et al. Effect of short-term cigarette smoke exposure on body weight, appetite and brain neuropeptide Y in mice. *Neuropsychopharmacology*. 2005;30:713-719.
6. Bush T, Lovejoy JC, Deprey M, Carpenter KM. The effect of tobacco cessation on weight gain, obesity, and diabetes risk. *Obesity (Silver Spring)*. 2016;24:1834-1841.
7. Carreras-Torres R, Johansson M, Haycock PC, et al. Role of obesity in smoking behaviour: Mendelian randomisation study in UK Biobank. *Bmj*. 2018;361:k1767.
8. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *Bmj*. 2018;362:k601.
9. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12:e1001779.
10. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203-209.
11. Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet*.

- 2019;51:237-244.
12. 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.
13. Buniello A, MacArthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 2019;47:D1005-d1012.
14. 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.
15. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 2017;13:e1007081.
16. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol.* 2016;40:597-608.
17. 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.
18. Park S, Lee S, Kim Y, et al. Kidney function and obstructive lung disease: a bidirectional Mendelian randomisation study. *Eur Respir J.* 2021;58.
19. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:7.
20. Zhao W, Ware EB, He Z, et al. Interaction between Social/Psychosocial Factors and Genetic Variants on Body Mass Index: A Gene-Environment Interaction Analysis in a Longitudinal Setting. *Int J Environ Res Public Health.* 2017;14.
21. Kvale MN, Hesselton S, Hoffmann TJ, et al. Genotyping Informatics and Quality

- Control for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics*. 2015;200:1051-1060.
22. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518:197-206.
 23. Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol*. 2018;33:947-952.
 24. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7.
 25. Dare S, Mackay DF, Pell JP. Relationship between smoking and obesity: a cross-sectional study of 499,504 middle-aged adults in the UK general population. *PLoS One*. 2015;10:e0123579.
 26. Winsløw UC, Rode L, Nordestgaard BG. High tobacco consumption lowers body weight: a Mendelian randomization study of the Copenhagen General Population Study. *Int J Epidemiol*. 2015;44:540-550.
 27. Dudbridge F. Commentary: Tobacco consumption and body weight: Mendelian randomization across a range of exposure. *Int J Epidemiol*. 2016;45:e1-e3.
 28. Peacock A, Leung J, Larney S, et al. Global statistics on alcohol, tobacco and illicit drug use: 2017 status report. *Addiction*. 2018;113:1905-1926.
 29. Degens H, Gayan-Ramirez G, van Hees HW. Smoking-induced skeletal muscle dysfunction: from evidence to mechanisms. *Am J Respir Crit Care Med*. 2015;191:620-625.
 30. Moslemi-Haghighi F, Rezaei I, Ghaffarinejad F, et al. Comparison of Physical Fitness among Smoker and Non-Smoker Men. *Addict Health*. 2011;3:15-19.
 31. Wang J, Bai Y, Zeng Z, et al. Association between life-course cigarette smoking and metabolic syndrome: a discovery-replication strategy. *Diabetol Metab Syndr*.

2022;14:11.

32. Yuan S, Larsson SC. A causal relationship between cigarette smoking and type 2 diabetes mellitus: A Mendelian randomization study. *Sci Rep*. 2019;9:19342.
33. Gkatzionis A, Burgess S. Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? *Int J Epidemiol*. 2019;48:691-701.

Table 1. Characteristics of the UK Biobank data.

Variables	Study sample (N = 337138)
Age (years)	58 [51, 63]
Sex	
Female	181027 (53.7%)
Male	156111 (46.3%)
Hypertension medication users	70021 (20.9%)
Dyslipidemia medication users	58533 (17.5%)
Diabetes mellitus	16179 (4.8%)
Systolic BP	136.5 [125.0,149.5]
Diastolic BP	82.0 [75.5,89.0]
Hb A1c (mmol/L)	35.1 [32.7,37.7]
Body mass index (kg/m ²)	26.72 [24.13,29.85]
Waist circumference (cm)	90 [80,99]
Hip circumference (cm)	102 [97, 108]
Waist-to-hip ratio	
Smoking history	
None-smoker	183640 (54.66%)
Ex-smoker	118403 (35.24%)
Current-smoker	33922 (10.10%)

Categorical variables are presented as N (%) and continuous variables as median [interquartile ranges].

Table 2. Causal estimates in the UK Biobank data.

Outcome	N of instrumented SNPs	MR-Egger intercept P value	MR methods	Causal estimates [beta (95% CI)]	P value
Body mass index (kg/m ²)	282	0.109	MR-IVW	0.28 (0.18, 0.38)	< 0.001
			Weighted median	0.29 (0.18, 0.4)	< 0.001
			Simple median	0.31 (0.2, 0.41)	< 0.001
			MR-Egger	0.23 (0.05, 0.42)	0.005
Waist circumference (cm)	292	0.116	MR-IVW	0.88 (0.66, 1.1)	< 0.001
			Weighted median	0.89 (0.64, 1.14)	< 0.001
			Simple median	0.89 (0.64, 1.14)	< 0.001
			MR-Egger	0.85 (0.41, 1.29)	< 0.001
Hip circumference (cm)	306	0.253	MR-IVW	0.4 (0.23, 0.57)	< 0.001
			Weighted median	0.4 (0.21, 0.6)	< 0.001
			Simple median	0.52 (0.34, 0.71)	< 0.001
			MR-Egger	0.26 (-0.06, 0.58)	0.06
Waist-to-hip ratio	299	0.523	MR-IVW	0.006 (0.005, 0.007)	< 0.001
			Weighted median	0.006 (0.004, 0.007)	< 0.001
			Simple median	0.006 (0.005, 0.007)	< 0.001
			MR-Egger	0.006 (0.004, 0.008)	< 0.001

MR = Mendelian randomization, SNP = single nucleotide polymorphism, MR-IVW = multiplicative random-effect inverse-variance weighted method

Causal estimates are scaled towards 2-fold increase in prevalence of tobacco smoking.

Table 3. Causal estimates in the GERA/GIANT meta-analysis.

Outcome	N of instrumented SNPs	MR-Egger intercept P value	MR methods	Causal estimates [beta (95% CI)]	P value
Body mass index (kg/m ²)	275	0.402	MR-IVW	0.03 (0.01, 0.05)	0.01
			Weighted median	0.03 (0, 0.05)	0.03
			Simple median	0.03 (0.01, 0.05)	0.02
			MR-Egger	0.04 (-0.01, 0.09)	0.045

MR = Mendelian randomization, SNP = single nucleotide polymorphism, MR-IVW = multiplicative random-effect inverse-variance weighted method

Causal estimates are scaled towards 2-fold increase in prevalence of tobacco smoking.

Figure legends

Figure 1. Study flow diagram. GWAS = genome-wide association study, QC = quality control, SNPs = single nucleotide polymorphism

Figure 2. Causal estimates from the Mendelian randomization analysis.

The x-axes show the causal estimates. The estimates by MR-IVW (multiplicative random-effect inverse variance weighted method), simple median, weighted median, and MR-Egger regression with bootstrapped standard error are presented. The outcome phenotypes assessed in the UK Biobank were body mass index, waist circumference, hip circumference, and waist-to-hip ratio. The replication analysis was performed for GERA/GIANT consortium data towards body mass index outcome. Causal estimates are scaled towards 2-fold increase in prevalence of tobacco smoking.

Genetic instrumental variable

Outcome dataset

GWAS meta-analysis for
tobacco smoking (ever smoker)
Including 1.2 million European
samples

Genetic instrument
for tobacco smoking
338 SNPs

Summary-level Mendelian randomization

Steiger filtering
Inverse-variance weighted
Pleiotropy-robust analysis

Primary outcome (UK Biobank)

Age 40-69 adults
at the UK biobank cohort
from 2006-2010
(N = 502,505)

White British ancestry
individuals passing sample QC
(N = 337,318)

Summary statistics for
body mass index,
waist circumference,
hip circumference,
and waist-to-hip ratio

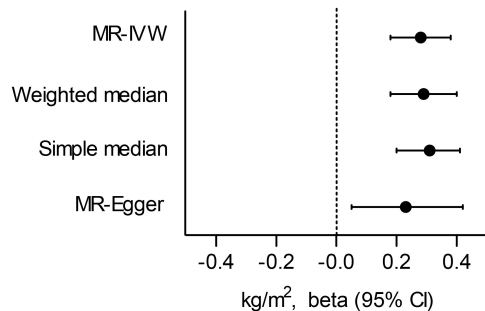
Replication (GERA/GIANT)

GWAS meta-analysis
For GERA/GIANT cohort data
(N = 364,487, > 90% European)

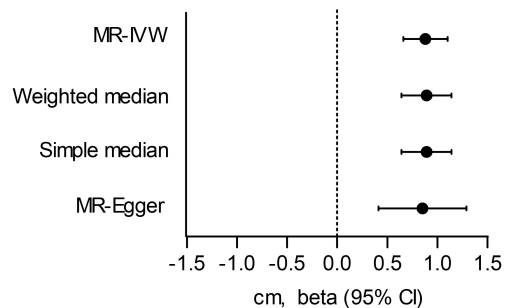
Summary statistics for
body mass index

UK Biobank data

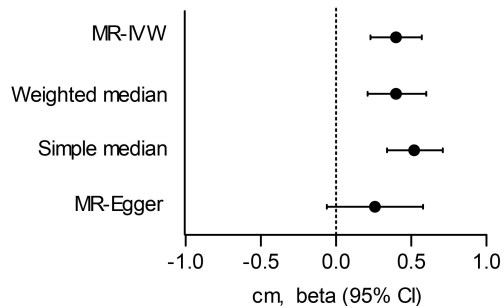
Body mass index



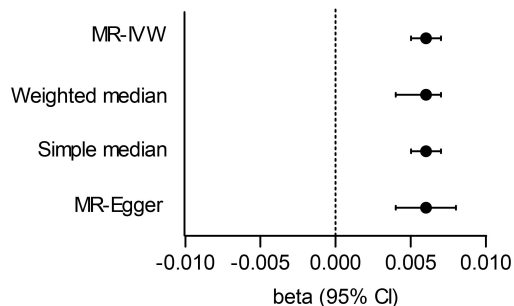
Waist circumference



Hip circumference



Waist-to-hip ratio



GERA/GIANT data

Body mass index

