Smoking and Dementia Status Among Older Americans: A Mendelian Randomization Analysis

Mingzhou Fu  
University of Michigan

Jessica D. Faul  
University of Michigan

Yuan Jin  
University of Michigan

Erin B. Ware  
University of Michigan

Kelly M. Bakulski  
bakulski@umich.edu  
University of Michigan

Research Article

Keywords: Smoking, polygenic score, dementia, cognitive impairment, epidemiology, Mendelian randomization

DOI: https://doi.org/10.21203/rs.3.rs-496150/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Smoking is associated with dementia, but causality has not been well-established. Given smoking has a strong genetic component, we used genetic predisposition to smoking, based on a genome-wide association study (GWAS) from the GWAS & Sequencing Consortium of Alcohol and Nicotine use, as an instrumental variable. In a Mendelian randomization framework, we assessed the relationships between smoking and impaired cognitive status using the 2010 wave of the Health and Retirement Study. Cognitive status was assessed using cognition tests and categorized in three levels (normal cognition, cognitive impairment-non dementia, dementia). Smoking status was self-reported and classified into never, former, and current smoking. We used multivariable logistic regressions, Wald-type ratio estimators, and two-sample Mendelian randomization methods to examine associations and infer causal relationships. Among European genetic ancestry participants (n = 8741), current smoking was associated with cognitive impairment (OR = 1.62, 95%CI: 1.29, 2.01) relative to normal cognition. Two-sample Mendelian randomization results showed an inferred causal effect of smoking initiation on late-onset Alzheimer’s type dementia (MR-Egger: OR = 1.021, 95%CI: 1.017, 1.025). We observed no association between current smoking and dementia, or between former smoking and any cognitive status. No associations between smoking status and cognitive status were observed in the African genetic ancestry sample (n = 2501). Smoking behavior can give rise to cognitive impairment; therefore, promotion of smoking cessation is also important for brain health. Studies on dose and duration of smoking on cognition are critically needed, as well as continued research in non-European genetic ancestry study samples.

Introduction

Dementia is a neurodegenerative disorder characterized by difficulties in a person’s daily life through memory loss, impaired language function, challenges in problem-solving, and changed cognitive status. In 2018 in the United States (US), 14% of people age over 70 were living with dementia, which cost an estimated $277 billion. Multiple factors contribute to the incidence of dementia, including genetic and environment factors. Cigarette smoking is one of the most prevalent but controversial dementia risk factors. With a prevalence of 20.6% among adults, smoking could be attributed to 10.8% of total dementia cases in the US. However, associations between cigarette smoking and dementia in many studies were inconsistent, and the causal relationship has not been rigorously assessed.

The likelihood of smoking is partially heritable. Based on meta-analysis, approximately 50% of the variation in smoking initiation and persistence was attributed to genetic factors. To date, genome-wide association studies (GWAS) have linked 566 genetic variants at 406 loci to smoking initiation, cessation, or heaviness. For example, a single nucleotide polymorphisms (SNP) of the 15q24 nicotinic acetylcholine receptor subunit alpha-5 (CHRNA5) gene (rs16969968[A]), is robustly associated with smoking behavior, and specifically nicotine dependence. Given the large genetic component to smoking, there is an opportunity to use smoking genetic predisposition as an epidemiologic tool.

Moving beyond association testing to rigorously assess causality is crucial for public health prevention. Mendelian randomization is an epidemiologic method using genetic variants as instrumental variables to assess inferred causality. It is analogous to a randomized control trial where individuals are randomized to carry genetic variants that may modify the risk of an exposure. Genetic variants are fixed at conception, which necessarily precedes the onset of exposures and health outcomes. Thus, Mendelian randomization can overcome many drawbacks of observational epidemiology studies, such as the potential for reverse causation, and it is useful when randomized trials would be unethical (as with smoking). Mendelian randomization has provided evidence for causal relationships between several modifiable risk factors and dementia, including glucose levels, insulin sensitivity, and cholesterol levels. Mendelian randomization studies have also provided evidence that body mass index is not causally associated with Alzheimer’s type dementia. Two studies have investigated causal effects of smoking on dementia or Alzheimer’s type dementia in participants of European ancestry and no causal effect was found using the individual genes, CHRNA5 or BDNF respectively, as instrumental variables. Mendelian randomization studies of smoking and dementia incorporating the broader scope of genetic predisposition to smoking are needed.

To investigate the association between smoking and dementia for causality, in the present study, we applied a Mendelian randomization framework. We first tested for replication of the reported association between smoking and cognitive status in our US nationally representative aging cohort study sample. We next evaluated whether polygenic predisposition to smoking was a valid instrumental variable for smoking status in two genetic ancestry groups. Then we deployed the two-sample Mendelian randomization to test an inferred causal relationship between smoking and dementia status.

Results

Sample descriptive statistics

The analytic sample included 11242 participants in the Health and Retirement Study (HRS) Wave 2010 with complete covariate information. Our analytic sample was younger, with lower proportions of never smokers and abnormal cognition compared to the excluded sample (n = 10786) (Supplementary Table 1). Table 1 shows characteristics of our primary analytic sample (European genetic ancestry n = 8741, African genetic ancestry n = 2501). Distributions of all covariates differed between the European and African genetic ancestry groups, except for proxy status on
Within the European genetic ancestry sample, age, sex, educational attainment, ever drinking alcohol, body mass index, depression score, and histories of stroke, hypertension, and diabetes were associated with both the outcome (cognitive status) and exposure (smoking status) and were treated as covariates in further adjusted models. Proxy respondent status, the number of impaired activities of daily living, and instrumental activities of daily living were also associated with cognitive impairment, though they may be a consequence of impaired cognition. Thus, we did not include those variables in our regression analyses (Table 2).

Assessment of genetic instrumental variable for smoking

Persons with higher polygenic scores for smoking were more likely to have smoked in the European ancestry sample (Table 3). A one standard deviation increase in smoking polygenic score was associated with $1.33 (95\% \text{CI}: 1.26, 1.40)$ times higher odds of former smoking and $1.52 (95\% \text{CI}: 1.41, 1.63)$ times higher odds of current smoking, relative to never smoking, after adjusting for age, sex, years of education and five ancestry-specific genetic principal components. With adjustment of smoking status, this association remained non-significant ($OR = 1.20, 95\% \text{CI}: 0.93, 1.11$). Together, these findings suggest the polygenic score for smoking is a valid genetic instrumental variable for smoking in relation to cognitive status.

Association between smoking status and cognitive status

Multivariable logistic regression analysis showed current smoking status was associated with cognitive impairment non-dementia in the European genetic ancestry sample (Table 5). In the primary adjusted model, current smokers relative to never smokers had $1.62 (95\% \text{CI}: 1.29, 2.02)$ times higher odds of cognitive impairment non-dementia relative to normal cognition. This association became attenuated after additional adjustment for health status ($OR = 1.43, 95\% \text{CI}: 1.12, 1.81$). According to population attributable fraction analysis, in our sample, $11.37\%$ of the cognitive impairment non-dementia cases were attributed to current smoking. No association was observed between current smoking and dementia, or between former smoking and cognitive status.

Using two sample Mendelian randomization analyses with GWAS summary statistics, we observed a causal relationship between smoking initiation and Alzheimer's type dementia (MR-Egger: $OR = 1.021, 95\% \text{CI}: 1.017, 1.025, n\text{SNP} = 7,193,484$). Thus, smoking initiation in the European genetic ancestry sample may be causally associated with $1.021$ times higher odds of Alzheimer's type dementia. Sensitivity analysis, we also performed a one-sample Mendelian randomization conduct within our HRS sample. However, the association between current smoking and cognitive impairment non-dementia remained non-significant ($OR = 1.02, 95\% \text{CI}: 0.93, 1.11$). This non-causal result was different from the association estimate (without any genetic factors) statistically ($p$ for heterogeneity = 0.48). Comparing with the two-sample Mendelian randomization results, the one-sample result showed a similar, weak, positive estimate, but with a wider confidence interval.

Sensitivity Analyses

In the European genetic ancestry sample, there was a weak but significant positive correlation between smoking polygenic score and Alzheimer's disease polygenic score ($r = 0.05, p < 0.01$), suggesting one or more pleiotropic genes affect smoking status and cognition simultaneously. To address this, we assessed using the primary smoking SNP (rs16969968) as an instrumental variable. The presence of the rs16969968[A] allele was not associated with smoking status, indicating the SNP was not a valid instrument for smoking status in this sample (Supplementary Table 2). As a negative control check, in the European subsample of never smokers only ($n = 3691$), smoking polygenic score was not associated with cognitive status (Supplementary Table 3). To assess the age specificity of our results we restricted analyses to participants between ages 60 and 90. Regression and Mendelian randomization results were similar in the older population to the full sample (Supplementary Table 4).

In the African genetic ancestry sample ($n = 2501$), we also found similar positive association between smoking polygenic score and smoking status. A one standard deviation increase in smoking polygenic score was associated with $1.22 (95\% \text{CI}: 1.09, 1.38)$ times higher odds of former smoking and $1.26 (95\% \text{CI}: 1.10, 1.45)$ times higher odds of current smoking, relative to never smoking, after adjusting for age, sex, years of education and five ancestry-specific genetic principal components. However, no associations were observed between smoking status and cognitive status in logistic regression analyses, so no further one-sample Mendelian randomization analyses were performed (Supplementary Table 5).
This Mendelian randomization analysis was conducted among older adults from the Health and Retirement Study in the 2010 wave. Consistent with prior results, in the European ancestry sample, after adjusting for age, sex, years of education, and five ancestry-specific genetic principal components, we observed a positive association between current smoking and cognitive impairment-non dementia (OR = 1.62, 95%CI: 1.29, 2.02). This is one of the first studies using a genome-wide polygenic score as an instrument (cumulative genetic risk for smoking) with a Mendelian randomization framework to examine the inferred causal relationship between smoking and cognitive status. Two sample Mendelian randomization results showed a causal relationship between smoking initiation and late-onset Alzheimer's type dementia (MR-Egger: OR = 1.021, 95%CI: 1.017, 1.025, nSNP = 7,193,484). These findings suggest current smoking may be causally associated with cognitive impairment.

Our multivariable regression association observed between current smoking and cognitive impairment-non dementia in the European ancestry sample was consistent with previous studies. Smokers may have elevated chronic oxidative stress in the brain and other organ systems, which may trigger the dementia-pathophysiological processes. Computed tomography and magnetic resonance based studies also found supportive evidence of abnormalities in brain morphology, perfusion and neurochemistry in smokers. Our null associations between former smoking and cognitive impairment were also similar to previous studies. Although there are several hypothesized pathways through which smoking may influence cognitive function, we were not able to identify specific mechanisms in this study.

Selective survival is a concern for both dementia and smoking. Smokers may die prematurely from other smoking-related diseases before developing dementia. To address this, we excluded participants older than age 90 and the oldest birth cohort, focusing our analyses on younger participants. Within the Health and Retirement Study, the genotyped respondents were longer-lived as compared with their non-genotyped respondents. The development of other smoking-related morbidity, such as cardiovascular diseases and cancer, may also impede participation in the longitudinal Health and Retirement Study. Thus, our sample might be biased toward healthier smokers – individuals who survived or did not...
experience significant smoking-related morbidity, and the smoking-dementia relationship was likely underestimated. This potential survival bias could also explain the null association between smoking and dementia found in our European sample (OR = 1.13, 95% CI: 0.66, 1.87).

Several limitations to this work are worth noting. First, although we attempted to account for confounding and population stratification, we are cautious about the findings. Future studies are warranted to replicate out findings in other cohorts and study populations. Second, cumulative genetic risk scores are subjective to weak instrument bias, meaning that if only one of the genetic variants is not a valid instrumental variable, the causal estimate will be biased and false positive error rates will be inflated. Third, cognitive status is not as well-defined as other more explicitly defined variables, such as smoking status, and was examined at only one time point, which almost certainly does not completely account for the lifetime variation in the trait instrumented by the genes. Next steps should include examining cognitive trajectories and dementia incidence. Results could also be strengthened by using more expansive smoking variables (such as number of cigarettes consumed or duration of smoking), and a longitudinal assessment of incident dementia.

Smoking behaviors are partially modifiable, which can be effectively altered, and promote a healthier aging population. Our study provided evidence to support a positive association between current smoking and cognitive impairment, and a causal relationship between smoking initiation and Alzheimer's type dementia. The null effect of former smoking suggests, while preliminary, a plausible protective effect of smoking cessation. Promotion of smoking cessation could be an effective strategy for lowering dementia prevalence over time at the population level.

In conclusion, in a cross-sectional analysis of the Health and Retirement Study, we observed current smoking was associated with cognitive impairment in older European ancestry adults, which is consistent with prior research. We conducted a two-sample Mendelian randomization analysis and observed an inferred causal relationship between smoking and impaired cognition. Interventions targeting contributing, confounding, or mediating pathways of this association may help attenuate the risk of cognitive impairment. Additional studies on dose and time-span effect of smoking behaviors on cognitive impairment are critically needed, as well as larger studies in diverse genetic ancestry study samples. Cessation of smoking is critical for public health more broadly, as well as cognitive health more specifically, shown here.

**Methods**

**Study population and design**

This cross-sectional study used the 2010 wave of the Health and Retirement Study, a publicly available national longitudinal panel study of individuals aged over 50 in the US. The Health and Retirement Study has collected data on health and economic information related to aging every two years since 1992 and more than 43,000 individuals have participated to date.

Detailed sample selection steps are shown in Supplementary Fig. 2. Since the risk factors and underlying neuropathological features of dementia are considerably different for people aged under 50 or over 90, we excluded participants who were in these age groups, as well as those in the Asset and Health Dynamics among the Oldest Old and Children of the Depression study, the two oldest cohorts of the Health and Retirement Study. Given the characteristics of progressive and irreversible cognition decline, we excluded individuals with reversed cognition, who were classified as normal cognition in 2010 but dementia in 2008, to minimize misclassification of cognitive status. To avoid potential recall bias, we also excluded those self-reported never smoking in 2010 but reported former or current smoking in 2008.

**Cognitive status assessment**

The main outcome variable in this analysis, cognitive status, was categorized in three levels as normal, cognitive impairment-non dementia, or dementia. The categorization method depended on whether the respondent could participate in the interview themselves or required a proxy respondent due to physical or cognitive problems. Self-respondent categories were based on a 27-point scale, according to participants’ performance on a series of cognition tests, including immediate and delayed 10-noun free recall, serial 7 subtraction, and backward count from 20. The proxy-respondent categories were based on an 11-point scale, using the proxy’s assessment of the respondent’s memory, whether the respondent had limitations in instrumental activities of daily living, and whether the respondent had difficulty completing the interview because of a cognitive limitation. Cognitive status cut points were established by Crimmins et al. and have been validated clinically and empirically with an area under curve score of 0.84. All cognition-related data were retrieved from the cross-wave imputation of cognitive functioning data provided by the Health and Retirement Study.

**Exposure assessment and covariates**

Smoking status was self-reported and classified into never, former, and current smoking. Smoking was defined as more than 100 cigarettes in a respondent’s lifetime (do not include pipes or cigars).

Other covariates used in our analysis included demographic characteristics, behavioral risk factors, and chronic health conditions. Age (years) was calculated by current wave year minus self-reported year of birth. Sex (male/female), years of education, proxy status (self/proxy-respondent), body mass index (kilograms/meters$^2$), alcohol consumption (ever drinking yes/no), and histories of hypertension, diabetes, and stroke (yes/no) were
self-reported. Center for Epidemiologic Studies of Depression (CESD) score of depressive symptoms was based on self-reported feelings in the past week (the higher the score, the more negative the respondent’s feelings). Activities of daily living (ADL) and instrumental activities of daily living (IADL) were self-proxy-reported according to proxy status, representing number of difficulties in performing two sets of tasks (ADL: bathing, eating, dressing, walking across a room, and getting in or out of bed; IADL: using a telephone, taking medication, handling money, shopping, and preparing meals). All variables were assessed at the 2010 wave and retrieved from the RAND Health and Retirement Study Longitudinal File.

### Genetic data

Respondents provided saliva samples after reading and signing a consent form during an enhanced face-to-face interview. Genotype measures were obtained using the Illumina HumanOmni2.5 BeadChip and genotyping was conducted by the Center for Inherited Disease Research. Genotype data that passed initial quality control were released to and analyzed by the Quality Assurance/Quality Control analysis team at the University of Washington. Details of the genotype collection and quality control are available elsewhere.

We used a polygenic score for smoking as the instrumental variable in our primary analysis. A polygenic score is a single quantitative measure of genetic risk, which aggregates multiple individual loci across the human genome and weights them by effect sizes derived from a GWAS. The smoking polygenic scores were created using SNPs and weights identified by a 2019 GWAS meta-analysis conducted by the GWAS & Sequencing Consortium of Alcohol and Nicotine use (GSCAN). The Alzheimer's disease polygenic scores used in sensitivity analysis were created using results from a 2019 GWAS conducted by Kunkle et al. excluding the APOE region. All the polygenic scores were standardized to a standard normal curve (mean = 0, standard deviation = 1) within genetic ancestry groups.

Genetic ancestry was identified through the union of self-reported race/ethnicity and principal component analysis on genome-wide SNPs calculated across all participants. Ancestry-specific genetic principal components were created in each ancestry sample by another principal component analysis to adjust for hidden population structure within ancestry. A binary variable of APOE ε4 allele carrier (having at least one copy of the ε4 allele, yes/no), and a categorical variable of smoking susceptibility rs16969968[A] carrier (0/1/2 copies) were derived from the imputed genetic data in the Health and Retirement Study which used a worldwide 1000 Genomes Project reference panel (phase I). All genetic data were downloaded from published datasets by the Health and Retirement Study.

### Statistical analyses

Analyses were carried out separately for European and African genetic ancestry groups. We used the χ² test and analysis of variance as appropriate to examine differences in baseline characteristics across included/excluded samples, different ancestry groups, exposure and outcome groups. Linear regression analysis was used to examine the associations of smoking polygenic score with other covariates.

An ordinal logistic regression was tried first to assess the proportional association between the three-level smoking and cognition variables. However, we observed a violation of the proportional odds assumption at all variables, except the smoking polygenic score, indicating heterogenous associations across different levels of smoking and cognitive status. Thus, we performed multivariable logistic regressions in four subsets of each ancestry sample, using never smoking and normal cognition as the reference groups. Primary models were adjusted for age, sex, years of education, and five ancestry-specific genetic principal components. Additional adjustment included dementia risk factors that were associated with smoking status. The results were presented as odds ratios (ORs) and 95% confidence interval (95%CIs). Population attributable fractions were also calculated for significant associations. P value < .05 was considered for statistical significance if not specified. Our primary analyses were in the European ancestry. The conceptual directed acyclic graph and study subsets are shown in Fig. 1.

Late-onset Alzheimer's disease is the most common form of dementia and may contribute to 60–70% of cases. Thus, we used late-onset Alzheimer's disease as the outcome phenotype for our two-sample Mendelian randomization study. We performed the two-sample Mendelian randomization analysis using the summary statistics from the same two GWASs (2019 GSCAN smoking and 2019 Kunkle Alzheimer's disease) that were used for polygenic score construction. We used the Mendelian randomization-Egger (MR-Egger) method, which implemented an adaption of Egger regression in Mendelian randomization study to allow for pleiotropic effects, with all overlapped SNPs in these two GWASs. We assumed that samples from these two GWASs were independent and no pruning of SNPs in linkage disequilibrium was performed. Mendelian Randomization results were also presented as ORs and 95%CIs. The two-sample Mendelian randomization was performed using the R package TwoSampleMR (version 0.5.6).

### Sensitivity analyses

We used the R package MendelianRandomization (version 0.4.2) to conduct an one-sample Mendelian randomization analysis, where the association of genetic variant with the exposure and with the outcome are estimated in the same sample, to examine whether smoking polygenic score was associated with cognitive status through its associations with smoking status. To evaluate the strength of the instrument, we used improvement χ² from the first-stage logistic regression, with values greater than 10 being taken as evidence against weak instruments. Wald-type ratio estimators were used to calculate the causal effect of smoking on cognitive status and the Delta method was used to estimate standard errors. To evaluate heterogeneity, we compared results from the Mendelian randomization analyses with standard logistic regression.
using interaction tests. Mendelian randomization methods are based on the following three assumptions: 1) relevance assumption: the genetic variants associate with the risk factor of interest; 2) independence assumption: there are no unmeasured confounders of the association between genetic variants and outcome; 3) exclusion restriction: the genetic variants affect the outcome only through the effect on the risk factor of interest. To test the robustness of the observed associations, we performed several sensitivity analyses to assess the potential violation of these assumptions. To check for pleiotropy, we tested for associations between the smoking polygenic score and general cognition polygenic score in our study sample. We also used a single SNP (rs16969968) as an instrumental variable instead to re-evaluate its associations with cognitive status. We further examined the association between smoking polygenic score and cognitive status in never smokers only, which served as a negative control.

The apolipoprotein E (APOE) e4 allele polymorphism is by far the strongest independent genetic risk factor for dementia. We added in APOE e4 allele carrier as a precision variable in our sensitivity analysis to reduce the standard errors of the regression models and hence shrink confidence intervals on coefficients of interest. Finally, we repeated our standard logistic regression analyses in an older population (age ≥60 and ≤90).

All analyses were conducted in R version 3.5.1. Code to produce all analyses in this manuscript are available (https://github.com/bakulskilab).

Declarations

Acknowledgments: We thank the participants of the Health and Retirement Study as well as the Health and Retirement Study staff for their data collection services.

Author Contributions: Conceptualization, E.B.W. and K.M.B.; methodology, M.F.; validation, Y.J.; formal analysis, M.F.; data curation, E.B.W. and M.F.; writing—original draft preparation, M.F.; writing—review and editing, K.M.B., E.B.W, and J.D.F.; supervision, E.B.W. and K.M.B.; funding acquisition, E.B.W. and K.M.B. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Institutional Review Board Statement: The Health and Retirement Study is sponsored by the National Institute on Aging (NIA U01AG009740) and is conducted by the University of Michigan, where written informed consent was approved by the Institutional Review Board. This secondary data analysis was exempt and not regulated, as determined by the Institutional Review Board at the University of Michigan (HUM00128220).

Informed Consent Statement: Informed consent was obtained from all individual participants included in the study.

Data Availability Statement: All data used in this study were publicly available and can be retrieved from the Health and Retirement Study website upon requests (https://hrs.isr.umich.edu/data-products/access-to-public-data). Code to produce all analyses in this manuscript are available (https://github.com/bakulskilab).

Funding: All authors were supported by grants from the National Institute on Aging (R01 AG055406, R01 AG067592). Dr. Bakulski was supported by grants from the National Institute for Environmental Health Sciences and the National Institute for Minority Health and Health Disparities (R01 ES025531; R01 ES025574; and R01 MD013299). Dr. Ware and Faul were both supported by a grant from the National Institute on Aging (R01 AG055654). Dr. Ware was additionally supported by a grant from the National Institute for Minority Health and Health Disparities (R01 MD012145). This work was supported by the University of Michigan, Michigan Lifestage Environmental Exposures and Disease (M-LEaD) National Institute for Environmental Health Sciences Core Center (P30 ES017885), the Michigan Alzheimer's Disease Center (P30 AG053760), and the Michigan Center for the Demography of Aging (P30 AG012846).

References

1. McKhann, G. M. et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement7, 263–269 (2011).
2. Facts and Figures. Alzheimer's Disease and Dementia https://alz.org/alzheimers-dementia/facts-figures.
3. Jellinger, K. A. The enigma of mixed dementia. Alzheimers Dement3, 40–53 (2007).
4. Jamal, A. et al. Current Cigarette Smoking Among Adults - United States, 2016. MMWR Morb. Mortal. Wkly. Rep.67, 53–59 (2018).
5. Barnes, D. E. & Yaffe, K. The projected effect of risk factor reduction on Alzheimer's disease prevalence. Lancet Neurol10, 819–828 (2011).
6. Lee, P. N. Smoking and Alzheimer's disease: a review of the epidemiological evidence. Neuroepidemiology13, 131–144 (1994).
7. Van Duijn, C. M. et al. Interaction between genetic and environmental risk factors for Alzheimer's disease: a reanalysis of case-control studies. Genet. Epidemiol11, 539–551 (1994).
8. Anstey, K. J., von Sanden, C., Salim, A. & O’Kearney, R. Smoking as a risk factor for dementia and cognitive decline: a meta-analysis of prospective studies. Am. J. Epidemiol166, 367–378 (2007).
9. Li, M. D., Cheng, R., Ma, J. Z. & Swan, G. E. A meta-analysis of estimated genetic and environmental effects on smoking behavior in male and female adult twins. *Addiction* **98**, 23–31 (2003).
10. Liu, M. *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.* **51**, 237–244 (2019).
11. Ware, J. J., van den Bree, M. B. M. & Munafò, M. R. Association of the CHRNA5-A3-B4 Gene Cluster With Heaviness of Smoking: A Meta-Analysis. *Nicotine Tob Res* **13**, 1167–1175 (2011).
12. Smith, G. D. & Ebrahim, S. ‘Mendelian randomization’ can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidem* **32**, 1–22 (2003).
13. Benn, M., Nordestgaard, B. G., Tybjærg-Hansen, A. & Frikke-Schmidt, R. Impact of glucose on risk of dementia: Mendelian randomisation studies in 115,875 individuals. *Diabetologia* **63**, 1151–1161 (2020).
14. Walter, S. *et al.* Diabetic Phenotypes and Late-Life Dementia Risk: A Mechanism-specific Mendelian Randomization Study. *Alzheimer Dis Assoc Disord* **30**, 15–20 (2016).
15. Zhang, X. *et al.* Causal association of circulating cholesterol levels with dementia: a mendelian randomization meta-analysis. *Transl Psychiatry* **10**, 145 (2020).
16. Mukherjee, S. *et al.* Genetically predicted body mass index and Alzheimer’s disease-related phenotypes in three large samples: Mendelian randomization analyses. *Alzheimers Dement* **11**, 1439–1451 (2015).
17. Nordestgaard, L. T., Tybjærg-Hansen, A., Nordestgaard, B. G. & Frikke-Schmidt, R. Body Mass Index and Risk of Alzheimer’s Disease: A Mendelian Randomization Study of 399,536 Individuals. *J Clin Endocrinol Metab* **102**, 2310–2320 (2017).
18. Kuzma, E. *et al.* Which Risk Factors Causally Influence Dementia? A Systematic Review of Mendelian Randomization Studies. *J Alzheimers Dis* **64**, 181–193.
19. Aggarwal, N. T. *et al.* The relation of cigarette smoking to incident Alzheimer’s disease in a biracial urban community population. *Neuroepidemiology* **26**, 140–146 (2006).
20. Juan, D. *et al.* A 2-year follow-up study of cigarette smoking and risk of dementia. *Eur. J. Neurol.* **11**, 277–282 (2004).
21. Ott, A. *et al.* Smoking and risk of dementia and Alzheimer’s disease in a population-based cohort study: the Rotterdam Study. *Lancet* **351**, 1840–1843 (1998).
22. Durazzo, T. C., Meyerhoff, D. J. & Nixon, S. J. Chronic cigarette smoking: implications for neurocognition and brain neurobiology. *Int J Environ Res Public Health* **7**, 3760–3791 (2010).
23. Reitz, C., den Heijer, T., van Duijn, C., Hofman, A. & Breteler, M. M. B. Relation between smoking and risk of dementia and Alzheimer disease: the Rotterdam Study. *Neurology* **69**, 998–1005 (2007).
24. Merchant, C. *et al.* The influence of smoking on the risk of Alzheimer’s disease. *Neurology* **52**, 1408–1412 (1999).
25. Ference, B. A., Majeed, F., Penumetcha, R., Flack, J. M. & Brook, R. D. Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or both: a 2 × 2 factorial Mendelian randomization study. *J Am Coll Cardiol* **65**, 1552–1561 (2015).
26. Jansen, H., Samani, N. J. & Schunkert, H. Mendelian randomization studies in coronary artery disease. *Eur Heart J* **35**, 1917–1924 (2014).
27. Davies, N. M., Holmes, M. V. & Davey Smith, G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
28. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* **23**, R89-98 (2014).
29. Martin, A. R. *et al.* Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
30. Domingue, B. W. *et al.* Mortality selection in a genetic sample and implications for association studies. *Int J Epidem* **46**, 1285–1294 (2017).
31. Burgess, S., Bowden, J., Fall, T., Ingelsson, E. & Thompson, S. G. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology* **28**, 30–42 (2017).
32. Ware, E. B. *et al.* Heterogeneity in polygenic scores for common human traits. *bioRxiv* 106062 (2017) doi:10.1101/106062.
33. Sonnega, A. *et al.* Cohort Profile: the Health and Retirement Study (HRS). *Int J Epidem* **43**, 576–585 (2014).
34. Bullain, S. S. & Corrada, M. M. Dementia in the oldest old. *Continuum (Minneap Minn)* **19**, 457–469 (2013).
35. Health and Retirement Study (HRS) & Asset and Health Dynamics among the Oldest Old (AHEAD) | sgim.org. https://www.sgim.org/communities/research/dataset-compendium/health-and-retirement-study-hrs–asset-and-health-dynamics-among-the-oldest-old-ahead.
36. Ofstedal, M. B. & Fisher, G. *Documentation of Cognitive Functioning Measures in the Health and Retirement Study*. http://hrsonline.isr.umich.edu/sitedocs/userg/dr-006.pdf (2005) doi:10.7826/ISR-UM.06.585031.001.05.0010.2005.
Tables

Table 1, Characteristics of study sample participants (n = 11242) stratified by genetic ancestry, Health and Retirement Study, Wave 2010.
| Categorical Variables [Count (Frequency)] | Overall | European ancestry | African ancestry | P-value<sup>b</sup> |
|-----------------------------------------|---------|-------------------|-----------------|-------------------|
| Smoking status at Wave 2010             | <0.001* |                   |                 |                   |
| Never                                   | 4659 (41.4%) | 3691 (42.2%) | 968 (38.7%) |                   |
| Former                                  | 4768 (42.4%) | 3823 (43.7%) | 945 (37.8%) |                   |
| Current                                 | 1815 (16.1%) | 1227 (14.0%) | 588 (23.5%) |                   |
| Cognitive Status at Wave 2010           | <0.001* |                   |                 |                   |
| Normal                                  | 9307 (82.8%) | 7654 (87.6%) | 1653 (66.1%) |                   |
| Cognitive Impairment-Non Dementia       | 1572 (14.0%) | 894 (10.2%) | 678 (27.1%) |                   |
| Dementia                                | 363 (3.23%) | 193 (2.21%) | 170 (6.80%) |                   |
| Sex (Female)                            | 6307 (56.1%) | 4802 (54.9%) | 1505 (60.2%) | <0.001* |
| Stroke history (Yes)                    | 718 (6.39%) | 502 (5.75%) | 216 (8.64%) | <0.001* |
| Hypertension history (Yes)              | 6492 (57.8%) | 4723 (54.1%) | 1769 (70.8%) | <0.001* |
| Diabetes history (Yes)                  | 2308 (20.5%) | 1597 (18.3%) | 711 (28.5%) | <0.001* |
| Drink status (Ever drinker)             | 6603 (58.7%) | 5342 (61.1%) | 1261 (50.4%) | <0.001* |
| Proxy status (Self-respondent)          | 11039 (98.2%) | 8573 (98.6%) | 2466 (98.6%) | 0.10 |
| APOE-ε4 allele carrier (Yes)            | 2678 (28.7%) | 1981 (26.4%) | 697 (38.0%) | <0.001* |
| Copies of rs16969968[A] allele          | <0.001* |                   |                 |                   |
| 0                                       | 4657 (52.9%) | 3084 (43.8%) | 1573 (89.3%) |                   |
| 1                                       | 3379 (38.4%) | 3197 (45.4%) | 182 (10.3%) |                   |
| 2                                       | 773 (8.78%) | 767 (10.9%) | 6 (0.34%) |                   |
| Continuous Variables [Mean (SD)]        |         |                   |                 |                   |
| Age (yrs)                               | 65.0 (8.93) | 65.8 (8.86) | 62.2 (8.62) | <0.001* |
| School years (yrs)                      | 13.2 (2.54) | 13.4 (2.44) | 12.4 (2.69) | <0.001* |
| Body Mass Index (kg/m<sup>2</sup>)      | 28.9 (6.22) | 28.4 (5.89) | 30.5 (7.03) | <0.001* |
| Activities of daily living              | 0.26 (0.80) | 0.22 (0.71) | 0.43 (1.01) | <0.001* |
| Instrumental activities of daily living  | 0.22 (0.72) | 0.19 (0.66) | 0.34 (0.86) | <0.001* |
| CESD score                              | 1.36 (1.93) | 1.22 (1.87) | 1.83 (2.07) | <0.001* |

Abbreviations: APOE, Apolipoprotein E; CESD, Center for Epidemiological Studies Depression; SD, Standard Deviation.

Notes:

[a] All the statistics including count, frequency, mean, SD, and P value were calculated based on non-missing data for each variable.

[b] The overall P value was calculated from chi-square test or analysis of variance for categorical or continuous variables as appropriate, interpreted as differences between groups.

Table 2. Bivariate characteristics stratified by cognitive or smoking status, Health and Retirement Study, Wave 2010, European ancestry sample (n = 8741)<sup>a</sup>
| Cognitive Status | Smoking Status |
|------------------|----------------|
| **Categorical Variables [Count (Frequency)]** | **Continuous Variables [Mean (SD)]** |
| Smoking status | **Overall** | **Normal** | **Impaired** | **Dementia** | **P-value** | **Never** | **Former** | **Current** | **P-value** |
| **n = 8741** | **n = 7654** | **n = 894** | **n = 193** | | | **n = 3691** | **n = 3823** | **n = 1227** | |
| Never | 3691 (42.2%) | 3315 (43.3%) | 311 (34.8%) | 65 (33.7%) | <0.001* | 2323 (62.9%) | 1803 (47.2%) | 676 (55.1%) | <0.001* |
| Former | 3823 (43.7%) | 3307 (43.2%) | 412 (46.1%) | 104 (53.9%) | | 176 (4.77%) | 245 (6.41%) | 91 (7.42%) | <0.001* |
| Current | 1227 (14.0%) | 1032 (13.5%) | 171 (19.1%) | 24 (12.4%) | | 1911 (51.8%) | 2217 (58.1%) | 595 (48.5%) | <0.001* |
| Sex (Female) | 4802 (54.9%) | 4325 (56.5%) | 395 (44.2%) | 82 (42.5%) | <0.001* | 650 (17.6%) | 759 (19.9%) | 188 (15.3%) | 0.001* |
| Stroke (Yes) | 512 (5.86%) | 344 (4.49%) | 110 (12.3%) | 58 (30.2%) | <0.001* | 176 (4.77%) | 245 (6.41%) | 91 (7.42%) | <0.001* |
| Hypertension (Yes) | 4723 (54.1%) | 4033 (52.8%) | 553 (61.9%) | 137 (71.4%) | <0.001* | 1911 (51.8%) | 2217 (58.1%) | 595 (48.5%) | <0.001* |
| Diabetes (Yes) | 1597 (18.3%) | 1319 (17.2%) | 230 (25.7%) | 48 (24.9%) | <0.001* | 650 (17.6%) | 759 (19.9%) | 188 (15.3%) | 0.001* |
| Drink status (Ever drinker) | 5342 (61.1%) | 4887 (63.8%) | 411 (46.0%) | 44 (22.8%) | <0.001* | 2074 (56.2%) | 2528 (66.1%) | 740 (60.3%) | <0.001* |
| Proxy status (Self-respondent) | 8573 (98.1%) | 7595 (99.2%) | 864 (96.6%) | 114 (59.1%) | <0.001* | 3637 (98.5%) | 3734 (97.7%) | 1202 (98.0%) | 0.02* |
| APOE-ε4 allele carrier (Yes) | 1981 (26.4%) | 1658 (25.4%) | 246 (30.9%) | 77 (43.0%) | <0.001* | 826 (25.9%) | 901 (27.0%) | 254 (26.0%) | 0.57 |
| Copies of rs16969968[A] allele | | | | | | | | | |
| 0 | 3084 (43.8%) | 2688 (43.8%) | 325 (43.4%) | 71 (42.3%) | | 1298 (43.5%) | 1371 (43.8%) | 415 (44.6%) | |
| 1 | 3197 (45.4%) | 2771 (45.2%) | 351 (46.9%) | 75 (44.6%) | | 1343 (45.0%) | 1436 (45.8%) | 418 (44.9%) | |
| 2 | 767 (10.9%) | 673 (11.0%) | 72 (9.63%) | 22 (13.1%) | | 344 (11.5%) | 326 (10.4%) | 97 (10.4%) | |

| Continuous Variables [Mean (SD)] | **Overall** | **Normal** | **Impaired** | **Dementia** |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| **n = 8741** | **n = 7654** | **n = 894** | **n = 193** |
| Age (yrs) | 65.8 (8.86) | 65.1 (8.67) | 70.2 (8.60) | 73.6 (7.77) | <0.001* | 65.5 (8.78) | 67.3 (8.71) | 62.1 (8.36) | <0.001* |
| School years (yrs) | 13.4 (2.44) | 13.7 (2.30) | 11.9 (2.65) | 11.6 (3.18) | <0.001* | 13.8 (2.40) | 13.4 (2.48) | 12.6 (2.23) | <0.001* |
| Body Mass Index (kg/m²) | 28.4 (5.89) | 28.5 (5.83) | 28.1 (6.32) | 26.1 (5.53) | <0.001* | 28.4 (5.87) | 29.8 (5.91) | 26.9 (5.65) | <0.001* |
| Activities of daily living | 0.22 (0.71) | 0.16 (0.59) | 0.40 (0.88) | 1.52 (1.90) | <0.001* | 0.17 (0.64) | 0.24 (0.75) | 0.28 (0.78) | <0.001* |
| Instrumental activities of daily living | 0.19 (0.66) | 0.12 (0.45) | 0.35 (0.81) | 2.24 (2.05) | <0.001* | 0.14 (0.59) | 0.21 (0.73) | 0.25 (0.66) | <0.001* |
| CESD score | 1.22 (1.87) | 1.14 (1.81) | 1.75 (2.13) | 2.50 (2.29) | <0.001* | 1.07 (1.71) | 1.16 (1.81) | 1.88 (2.34) | <0.001* |
| Smoking polygenic risk score | 0.01 (0.99) | 0.00 (0.99) | 0.06 (1.03) | 0.14 (1.00) | 0.03* | -0.16 (0.97) | 0.08 (0.98) | 0.25 (1.03) | <0.001* |
| AD polygenic risk score | 0.01 (1.00) | -0.01 (1.00) | 0.14 (1.03) | 0.02 (1.04) | 0.001* | 0.01 (1.02) | 0.00 (1.00) | 0.07 (0.98) | 0.17 |
**Abbreviations:** AD, Alzheimer’s disease; APOE, Apolipoprotein E; CESD, Center for Epidemiological Studies Depression; SD, Standard Deviation

**Notes:**

[a] All the statistics including count, frequency, mean, SD, and P value were calculated based on non-missing data for each variable.

[b] The overall P value was calculated from chi-square test or analysis of variance for categorical or continuous variables as appropriate, interpreted as differences between groups.

**Table 3.** Associations between polygenic risk score for smoking and smoking status, Health and Retirement Study, Wave 2010, European ancestry sample (n = 8741)\(^a\)

|                | Impaired Cognition & Normal Sample | Dementia & Normal Sample | Overall Sample |
|----------------|-----------------------------------|--------------------------|----------------|
|                | N  | OR (95 CI%) | Improvement C\(^2\)\(^b\) | N  | OR (95 CI%) | Improvement C\(^2\) | N  | OR (95 CI%) | Improvement C\(^2\) |
| **Former vs. Never** |    |            |                        |    |            |                        |    |            |                        |
| **Crude**      | 7345 | 1.28 (1.22, 1.34) | -                       | 6791 | 1.28 (1.21, 1.34) | -                       | 7514 | 1.29 (1.23, 1.35) | -                       |
| **Adjusted\(^c\)** | 7345 | 1.32 (1.26, 1.39) | 118.1                   | 6791 | 1.32 (1.25, 1.39) | 109.2                   | 7514 | 1.33 (1.26, 1.39) | 126.2                   |
| **Current vs. Never** |    |            |                        |    |            |                        |    |            |                        |
| **Crude**      | 4829 | 1.51 (1.41, 1.62) | -                       | 4436 | 1.52 (1.41, 1.63) | -                       | 4918 | 1.52 (1.42, 1.63) | -                       |
| **Adjusted**   | 4829 | 1.51 (1.41, 1.63) | 123.6                   | 4436 | 1.61 (1.40, 1.64) | 107.0                   | 4918 | 1.52 (1.41, 1.63) | 127.0                   |

**Abbreviations:** CI, Confidence Interval; OR, Odds Ratio

**Notes:**

[a] All the values were based on results from standard logistic regression analyses in each sample, in which “never smoking” was used as the reference group.

[b] Calculated by 2*(log likelihood of full model - log likelihood of reduced model). Statistics larger than 10 indicates a valid instrument in convention.

[c] Adjusted for age, sex, education and five ancestry-specific principal component sets.

**Table 4.** Associations between polygenic score for smoking and cognitive status, Health and Retirement Study, Wave 2010, European ancestry sample (n = 8741)\(^a\)
|                          | Impaired Cognition vs. Normal | Dementia vs. Normal |
|--------------------------|------------------------------|---------------------|
|                          | Former vs. Never              | Current vs. Never    | Former vs. Never | Current vs. Never |
|                          | N    | OR (95 CI%)       | N    | OR (95 CI%)       | N    | OR (95 CI%)       | N    | OR (95 CI%)       |
| **Total effect of smoking polygenic score** |                  |                    |                  |                  |                  |                  |                  |
| Crude                    | 7345 | 1.06 (0.98, 1.14) | 4829 | 1.05 (0.95, 1.15) | 6791 | 1.16 (0.99, 1.36) | 4436 | 1.02 (0.83, 1.27) |
| Adjusted (demographic)   | 7345 | 1.02 (0.93, 1.11) | 4829 | 1.03 (0.93, 1.14) | 6791 | 1.11 (0.94, 1.32) | 4436 | 1.00 (0.80, 1.26) |
| Adjusted (health status) | 7191 | 1.01 (0.93, 1.11) | 4723 | 1.02 (0.91, 1.13) | 6590 | 1.04 (0.83, 1.31) | 4326 | 0.97 (0.73, 1.28) |
| Adjusted (AD genetics)   | 5856 | 1.02 (0.92, 1.12) | 3772 | 1.03 (0.91, 1.16) | 5353 | 1.00 (0.78, 1.29) | 3440 | 0.96 (0.69, 1.33) |
| **Direct effect of smoking polygenic score (adjusting for smoking status)** |                  |                    |                  |                  |                  |                  |                  |
| Crude                    | 7345 | 1.04 (0.96, 1.13) | 4829 | 1.00 (0.91, 1.10) | 6791 | 1.13 (0.97, 1.32) | 4436 | 1.01 (0.82, 1.25) |
| Adjusted (demographic)   | 7345 | 1.02 (0.93, 1.11) | 4829 | 0.99 (0.89, 1.10) | 6791 | 1.11 (0.93, 1.31) | 4436 | 0.99 (0.79, 1.25) |
| Adjusted (health status) | 7191 | 1.01 (0.92, 1.11) | 4723 | 0.99 (0.89, 1.10) | 6590 | 1.05 (0.84, 1.33) | 4326 | 0.98 (0.73, 1.30) |
| Adjusted (AD genetics)   | 5856 | 1.02 (0.92, 1.12) | 3772 | 1.00 (0.88, 1.12) | 5353 | 1.01 (0.78, 1.30) | 3440 | 0.97 (0.70, 1.36) |

**Abbreviations:** AD, Alzheimer’s disease; APOE, Apolipoprotein E; CESD, Center for Epidemiological Studies Depression; CI, Confidence Interval; OR, Odds Ratio; PGS, polygenic risk score

**Notes:**

[a] All the values were based on results from multivariable logistic regression in each sample, in which “never smoking” and “normal cognitive status” were used as the reference groups.

[b] Adjusted for age, sex, years of education and five ancestry-specific principal component sets.

[c] Adjusted for ever drink alcohol, history of stroke, hypertension, diabetes, BMI, and CESD score in addition to variables in [b].

[d] Adjusted for APOE4 allele status and AD PGS in addition to variables in [c].

**Table 5.** Associations between smoking status and cognitive status, Health and Retirement Study, Wave 2010, European ancestry sample"
### Former vs. Never

|                          | Logistic Regression | Wald-type/ratio | \(P\) for heterogeneity<sup>b</sup> |                          | Logistic Regression | Wald-type/ratio | \(P\) for heterogeneity<sup>b</sup> |
|--------------------------|---------------------|----------------|------------------------------------|--------------------------|---------------------|----------------|------------------------------------|
| **Impaired vs. Normal**  |                     |                |                                    |                          |                     |                |                                    |
| Crude                    | 7345                | 1.33           | (1.14, 1.55)                       | 1.26                     | (0.92, 1.73)       | 0.76           | 4829                 | 1.77           | (1.44, 2.15)                       | 1.12           | (0.89, 1.41)                       | 0.003*        |
| Adjusted (demographic)<sup>c</sup> | 7345                | 0.98           | (0.83, 1.16)                       | 1.06                     | (0.78, 1.45)       | 0.63           | 4829                 | 1.62           | (1.29, 2.02)                       | 1.07           | (0.83, 1.37)                       | 0.01*        |
| Adjusted (health status)<sup>d</sup> | 7191                | 1.02           | (0.86, 1.22)                       | 1.05                     | (0.76, 1.44)       | 0.88           | 4723                 | 1.43           | (1.12, 1.81)                       | 1.04           | (0.80, 1.35)                       | 0.06         |
| Adjusted (AD genetics)<sup>e</sup> | 5856                | 1.03           | (0.85, 1.25)                       | 1.07                     | (0.75, 1.52)       | 0.86           | 3772                 | 1.44           | (1.11, 1.87)                       | 1.07           | (0.80, 1.42)                       | 0.10         |
| **Dementia vs. Normal**  |                     |                |                                    |                          |                     |                |                                    |
| Crude                    | 6791                | 1.60           | (1.18, 2.20)                       | 1.84                     | (0.97, 3.48)       | 0.70           | 4436                 | 1.19           | (0.73, 1.88)                       | 1.06           | (0.64, 1.76)                       | 0.74         |
| Adjusted (demographic)   | 6791                | 1.13           | (0.81, 1.58)                       | 1.47                     | (0.80, 2.69)       | 0.45           | 4436                 | 1.13           | (0.66, 1.87)                       | 1.01           | (0.58, 1.75)                       | 0.76         |
| Adjusted (health status) | 6590                | 0.90           | (0.57, 1.42)                       | 1.16                     | (0.51, 2.64)       | 0.59           | 4326                 | 0.90           | (0.46, 1.70)                       | 0.92           | (0.47, 1.81)                       | 0.94         |
| Adjusted (AD genetics)   | 5353                | 0.94           | (0.57, 1.55)                       | 1.01                     | (0.41, 2.52)       | 0.90           | 3440                 | 0.82           | (0.37, 1.70)                       | 0.90           | (0.41, 1.98)                       | 0.84         |

**Abbreviations:** CESD, Center for Epidemiological Studies Depression; CI, Confidence Interval; MR, Mendelian Randomization; OR, Odds Ratio; AD, Alzheimer's disease

**Notes:**

[a] All values were based on results from multivariable logistic regression or Mendelian randomization analyses in each sample, in which "normal cognitive status" and "never smoking" were used as the reference groups.

[b] \(P\) value represents the statistical significance of the test of heterogeneity between logistic regression and Mendelian randomization.

[c] Adjusted for age, sex, years of education, and five ancestry-specific principal component sets.

[d] Adjusted for ever drink alcohol, history of stroke, hypertension, diabetes, BMI, and CESD score in addition to variables in [c].

[e] Adjusted for APOE4 allele status and AD PGS in addition to variables in [d].

**Figures**
Figure 1

Mendelian randomization analyses structure and subsets of study population, Health and Retirement Study, Wave 2010 (n = 11242)\(^a\) Notes: [a] Mendelian randomization assumptions: (i) the genetic variant must be associated with the exposure, (ii) the genetic variant must not be associated with any confounder of the exposure-outcome association and (iii) the genetic variant must be associated with the outcome only via the exposure. [b] Subsets used in standard logistic regression and Mendelian randomization analyses. All primary models adjusted for age, sex, years of education, and five ancestry-specific genetic principal components.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SmokingMRSupmzfu.docx