medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Initiation of smoking and susceptibility to type 2 diabetes: a mendelian randomization study

Christopher S Thom MD, PhD, Zhuoran Ding, Benjamin F Voight, PhD

Christopher S Thom, MD, PhD,
Division of Neonatology, Children's Hospital of Philadelphia, Philadelphia, PA, USA, 19104

Zhuoran Ding, Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA, 19104

Benjamin F Voight, PhD, Department of Systems Pharmacology and Translational Therapeutics, Department of Genetics, and Institute of Translational Medicine and Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA, 19104

Corresponding author:
Benjamin F Voight
University of Pennsylvania Perelman School of Medicine
3400 Civic Center Blvd
10-126 Smilow Center for Translational Research
Philadelphia, PA 19104
bvoight@pennmedicine.upenn.edu
Phone: 215-746-8083

Running Title: Smoking and type 2 diabetes risk.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Objectives

To evaluate the association between smoking behavior and type 2 diabetes risk using genetic data.

Research Design and Methods

We obtained summary data from genome-wide association studies (GWAS) related to smoking behavior (n=1.2 million people) and type 2 diabetes (n=898,130) in individuals of European ancestry. Mendelian randomization analysis was performed using these summary-level statistics. Specifically, we assessed if genetically determined smoking initiation, smoking cessation, and smoking heaviness (number of cigarettes per day) modulated susceptibility to type 2 diabetes.

Results

We observed that for each 2-fold increased odds of initiating smoking behavior corresponded with a 21% increased susceptibility to type 2 diabetes (OR = 1.21, 95% CI: 1.18-1.24, $P=1x10^{-12}$). Our effect was consistent across instrumental variable methodological approaches, and we did not observe evidence of systematic biases to our effect estimate. When analyzed using an analogous approach, we found that the effect of smoking initiation increasing diabetes risk was equivalent to the effect of smoking behavior on increase coronary artery disease, an established causal risk factor for cardiovascular disease.

Conclusions

Genetic data supports the hypothesis that smoking initiation corresponds with increased risk for type 2 diabetes. These findings are consistent with previous clinical epidemiological association between smoking behavior and elevated type 2 diabetes risk. Individual-level MR studies are necessary to estimate mediators and confounders related to this association.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Main Text

Introduction

Type 2 diabetes is a leading causes of death worldwide (1). Type 2 diabetes and related comorbidities are increasing global epidemics, expected to place ever larger demands on health care systems. As such, risk factors that cause type 2 diabetes that can be readily intervened upon are crucial for prevention to public health.

Recent clinical and epidemiologic studies have demonstrated association with smoking with increased diabetes risk (2–6), suggesting that smoking might be a modifiable risk factor for type 2 diabetes. Smoking can perturb glycemic regulation (5,7), and murine studies recently revealed a common biological pathway underlying nicotine addiction and type 2 diabetes (8). Although smoking is associated with perturbations in a number of cardiometabolic traits (9–11), to our knowledge a causal genetic association between smoking and type 2 diabetes in humans has not yet been established.

Recent collections of genetic data for smoking behavior and diabetes risk now provide a vehicle through which evidence of causality can be evaluated, using the framework of mendelian randomization (MR). Formally, MR uses genetic variants associated with an exposure of interest (here, smoking behavior) to create an instrumental variable to estimate a potential causal effect of genetically-determined exposure on an outcome (here, type 2 diabetes). Because alleles are randomly allocated at meiosis, and due to the fact that genotype precedes phenotype, this approach with care can address issues of confounding and reverse causality that prevent inference in prospective or cross-sectional cohort studies (12). Using the largest genome wide association studies (GWAS) for both smoking behavior (13) and type 2 diabetes (14) reported to date, we applied the framework of MR to demonstrate evidence supporting smoking initiation as a risk factor for increasing susceptibility to type 2 diabetes.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Research Design and Methods

Genome-wide summary statistics collection
We analyzed publicly available GWAS summary statistics for smoking initiation (n=1.2 million individuals), smoking cessation (n=547,219), and cigarettes per day (n=337,334) (13), as well as type 2 diabetes (n=898,130) (14), and coronary artery disease (CAD, n=~546,000) (15). We used data from individuals of European ancestry only. All data sets were analyzed in genome build hg19.

Genetic variant selection related to smoking exposures
We created genetic instruments related to smoking traits, including 'smoking initiation', 'smoking cessation' and 'cigarettes per day' (smoking heaviness) (13). For each smoking trait, we used PLINK (16) to identify single nucleotide polymorphisms within independent linkage disequilibrium blocks (EUR r^2<0.01) in 250kb regions. Analogous instrumental variable (IV) results were obtained different LD pruning thresholds (i.e., r^2<0.1 and r^2<0.001, data not shown). After harmonization, full IVs included 341 single nucleotide variants (SNPs) for smoking initiation (**Supplemental Table 1**), 18 SNPs for smoking cessation (**Supplemental Table 2**), and 42 SNPs for cigarettes per day (**Supplemental Table 3**) on type 2 diabetes. Analogous workflows generated IVs for smoking traits on CAD (n= 342 SNPs, **Supplemental Table 4**).

We used mRnd (http://cnsgenomics.com/shiny/mRnd/, (17)) to estimate the strength F-statistics of these IVs (**Supplemental Table 5**). Smoking initiation or cessation were input as binary exposure variables, and smoking heaviness as a continuous exposure variable. We calculated proportion of genetic inheritance explained in each exposure variable per Shim *et. al.* (18). F-statistics greater than 10 were presumed to be devoid of weak instrument bias.

Mendelian randomization and causal effect estimation
We performed two-sample MR (TwoSample MR package v0.4.25 (12)) using R (v3.6.1). We present causal estimates from inverse variance weighted (random effects model), weighted median, and MR Egger regression methods. We analyzed for pleiotropic bias using Egger regression intercepts, wherein significant non-zero intercepts can imply directional bias among IVs (19).

Genetic correlation estimates
Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) (20). Summary statistics were munged and analyzed for each trait. Presented data reflecting genetic correlation estimates (i.e., r_g values).

Statistical analysis
Effects from smoking exposure on type 2 diabetes were analyzed using inverse variance weighted, weighted median, and MR-Egger regression measures. Because Cochran's Q test (included in the TwoSample MR package v0.4.25 (12)) found heterogeneity our IVs, we utilized the random-effect model when performing inverse-variance weighted MR. We also performed a sensitivity analysis using IV pruning via MR-PRESSO (21) for the single trait (smoking initiation). The resultant IV included 325 SNPs from an initial 341, but MR results were not qualitatively different than the full instrument (see Supplementary Material), and thus we focus our presentation on MR results using all SNPs. Statistical significance was defined as $P < 0.05$.

We converted causal effect estimates into odds ratios using previously described methods (22). By multiplying causal effect estimates by ln(2) and taking the exponentiated value (=exp^[ln(2)*effect]), we calculated the change in type 2 diabetes odds ratio that corresponded

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

with a 2-fold change in smoking initiation or smoking cessation risk. For smoking heaviness, a continuous variable, we exponentiated the causal effect estimate (=exp^[effect]) to reach a value reflecting the change in odds ratio per unit increase in exposure.

Coding scripts and data sets

All relevant coding scripts and data sets can be found on Github (https://github.com/thomchr/SmkT2D). All data and coding scripts are also available upon request.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Results

Increased odds of smoking initiation elevated risk for type 2 diabetes

To estimate a causal effect between odds of smoking initiation and susceptibility to type 2 diabetes, we performed causal inference analysis using the framework of mendelian randomization. Using summary statistics obtained from genome-wide association studies of these traits, we generated an instrumental variable (IV) for smoking initiation (13), applied to data for type 2 diabetes (14). Our genetic instrument comprised 341 linkage-independent single nucleotide polymorphisms (SNPs, EUR r^2<0.01) related to this smoking behavior. Our genetic instrument was not subject to weak-instrument bias (F-statistic = 4926). We observed that both inverse-variance weighted ($P=1.0x10^{-12}$) and median-weighted ($P=6.9x10^{-12}$) MR methods demonstrated positive association between genetically determined increase in odds of smoking initiation and increased susceptibility to type 2 diabetes (**Fig. 1**). As a sensitivity analysis, we performed the MR-Egger regression test, and found no positive evidence of systematic bias in our estimated effect (MR-Egger intercept term P=0.87). These data support the hypothesis that increased odds of smoking initiation corresponds with an increase of type 2 diabetes; each 2-fold increase in smoking initiation corresponding to a 21% increased type 2 diabetes risk (odds ratio, OR=1.21, 95% confidence interval = 1.18-1.24, by inverse variance weighted method).

Increased risk of type 2 diabetes by smoking initiation is equivalent to that of coronary artery disease

To contrast our estimated effect to established causal relationships, we next performed an equivalent analysis using our smoking initiation instrumental variable on susceptibility to coronary artery disease (CAD). As with the above causal inference experiment with type 2 diabetes, our genetic instrument was devoid of weak instrument bias (F-statistic = 3780). We observed that a 2-fold increase in smoking initiation increased odds of coronary artery disease risk by 21% (OR = 1.21, 95% CI = 1.19-1.24 by inverse variance weighted method). The positive estimated causal effect on CAD agreed across MR methods (**Fig. 1**), and was consistent with previous reports (6). Strikingly, the estimated effect of smoking initiation on CAD risk was nearly identical to increased type 2 diabetes risk from this behavior (**Fig. 1**). These results indicate that the impact of smoking initiation on type 2 diabetes risk may have the same relative effect as this behavior on risk to CAD.

Smoking cessation and heaviness were not associated with type 2 diabetes risk

We next generated instrumental variables for additional smoking behaviors, including smoking cessation and smoking heaviness (cigarettes per day). The number of instruments for both of these traits were substantially smaller (n=18 and n=42 SNPs, respectively). F-statistics suggested that these instruments had no weak-instrument bias (F-statistic = 1737 for smoking cessation, F-statistic = 9897 for smoking heaviness). However, we did not observe evidence of association with either of these traits and type 2 diabetes risk (**Supplemental Fig. 1** and **Supplemental Table 6**). A lack of association for these smoking traits with type 2 diabetes risk may be driven by lack of statistical power to detect relatively small effect sizes.

Smoking traits are genetically correlated with type 2 diabetes risk

If the results from the above were due to a lack of statistical power, we might expect to observe positive genetic correlation between smoking behaviors and type 2 diabetes risk. Therefore, we quantified the extent to which genetic susceptibility to these three smoking traits correlate with

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

genetic susceptibility to type 2 diabetes. Using linkage disequilibrium score regression (LDSC) (20), we observed significant and positive genetic correlations between smoking traits and type 2 diabetes (**Table 1**).These results suggests that future genetic studies that explain smoking behavior and type 2 diabetes might clarify the effects of these smoking traits on type 2 diabetes.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Conclusions

Despite clinical observations (6), causal association between smoking and type 2 diabetes has not been established previously. Using the largest available repositories of meta-analyzed genetic summary statistics (13,14), we provide to our knowledge the first report of genetic evidence supporting increased risk of smoking initiation as a risk factor increased type 2 diabetes. The directionality of this finding is statistically and clinically significant, and concordant with the effects of smoking on CAD risk (**Fig. 1**). Although interpretation of these findings is limited to individuals of European ancestry, we expect that well-powered multi-ethnic GWAS will ultimately confirm this result in other populations.

These findings confirm and extend clinical observations relating smoking to increased type 2 diabetes risk, adding to numerous reports regarding negative effects of smoking on cardiometabolic health identified by medical practitioners and policymakers. A recent report detailed one mechanism linking smoking with type 2 diabetes risk (8). We anticipate that future validation research will establish clinically relevant biological mechanisms related to other loci as well. Although the effects of smoking may be mediated in part by comorbid cardiometabolic conditions, future multivariate MR studies will be necessary to investigate these associations. Importantly, these results show for the first time a genetic association between smoking and increased type 2 diabetes risk.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Acknowledgements

CST and BFV designed the study. CST, ZD, and BFV conducted, analyzed, and/or interpreted experimental data. CST and BFV wrote the paper. BFV acts as guarantor and corresponding author for this manuscript. The authors declare no conflicts of interest.

This work was supported by R56DK101478 and R01HG010067 (BFV), a Linda Pechenik Montague Investigator Award (BFV), T32HD043021 (CST), and an American Academy of Pediatrics Marshall Klaus Neonatal-Perinatal Research Award (CST). All authors confirm independence from funders and that all authors had full access to all of the statistics in the study. As such they can attest to the integrity of the data and the accuracy of the data analysis.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. Willi C, Bodenmann P, Ghali WA, Willi C. Active smoking and the risk of type 2 diabetes—a systematic review and meta-analysis. JAMA [Internet]. 2010 Dec 12 [cited 2019 Sep 10];298(22):2654. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.298.22.2654
2. Pan A, Wang Y, Talaei M, Hu FB, Wu T. Relation of active, passive, and quitting smoking with incident type 2 diabetes: A systematic review and meta-analysis. Lancet Diabetes Endocrinol [Internet]. 2015;3(12):958–67. Available from: http://dx.doi.org/10.1016/S2213-8587(15)00316-2
3. Ronnemaa T, Rônnemaa EM, Puukka P, Pyorälä MK, Laakso M. Smoking is independently associated with high plasma insulin levels in nondiabetic men. Diabetes Care [Internet]. 1996 Nov 1 [cited 2019 Sep 10];19(11):1229–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8908385
4. Davey Smith G, Bracha Y, Svendsen KH, Neaton JD, Haffner SM, Kuller LH. Incidence of Type 2 Diabetes in the Randomized Multiple Risk Factor Intervention Trial. Ann Intern Med [Internet]. 2005 Mar 1 [cited 2019 Sep 10];142(5):313. Available from: http://annals.org/article.aspx?doi=10.7326/0003-4819-142-5-200503010-00006
5. Sargeant LA, Khaw KT, Bingham S, Day NE, Luben RN, Oakes S, et al. Cigarette smoking and glycaemia: The EPIC-Norfolk study. Int J Epidemiol [Internet]. 2001 Jun 1 [cited 2019 Sep 10];30(3):547–54. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/30.3.547
6. Carter AR, Gill D, Davies NM, Taylor AE, Tillmann T, Vaucher J, et al. Understanding the consequences of education inequality on cardiovascular disease: Mendelian randomisation study. BMJ [Internet]. 2019 May 22 [cited 2019 Nov 1];365:1–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31122926
7. Ding R, Huang T, Han J. Diet/lifestyle and risk of diabetes and glycemic traits: A Mendelian randomization study. Lipids Health Dis [Internet]. 2018 Jan 29 [cited 2019 Sep 10];17(1):18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29375034
8. Duncan A, Heyer MP, Ishikawa M, Caligiuri SPB, Liu X, Chen Z, et al. Habenular TCF7L2 links nicotine addiction to diabetes. Nature [Internet]. 2019 Oct 16 [cited 2019 Oct 17];574(7778):372–7. Available from: http://www.nature.com/articles/s41586-019-1653-x
9. Carreras-Torres R, Johansson M, Haycock PC, Relton CL, Davey Smith G, Brennan P, et al. Role of obesity in smoking behaviour: Mendelian randomisation study in UK Biobank. BMJ [Internet]. 2018 May 16 [cited 2019 Sep 10];361:k1767. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29769355
10. Åsvold BO, Bjørngaard JH, Carslake D, Gabrielsen ME, Skorpen F, Smith GD, et al. Causal associations of tobacco smoking with cardiovascular risk factors: A Mendelian randomization analysis of the HUNT study in Norway. Int J Epidemiol [Internet]. 2014 Oct 1 [cited 2019 Sep 10];43(5):1458–70. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyu113
11. Thorgeirsson TE, Gudbjartsson DF, Sulem P, Besenbacher S, Styrkarsdottir U, Thorleifsson G, et al. A common biological basis of obesity and nicotine addiction. Transl Psychiatry [Internet]. 2013 Oct 1 [cited 2019 Sep 10];3(10):1–7. Available from: http://www.nature.com/articles/tp201381
12. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-base platform supports systematic causal inference across the human phenome. Elife [Internet]. 2018 [cited 2019 Oct 18];7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29846171
13. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Nat Genet [Internet]. 2019 [cited 2019 Oct 18];51(2):237–44. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6358542/

14. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet. 2018;50(11):1505–13.
15. Van Der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. Circ Res [Internet]. 2018 [cited 2019 Oct 24];122(3):433–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29212778
16. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience [Internet]. 2015 Dec 25 [cited 2019 Oct 18];4(1):7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25722852
17. Brion M-JA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol [Internet]. 2013 Oct 1 [cited 2019 Nov 20];42(5):1497–501. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyt179
18. Shim H, Chasman DI, Smith JD, Mora S, Ridker PM, Nickerson DA, et al. A Multivariate Genome-Wide Association Analysis of 10 LDL Subfractions, and Their Response to Statin Treatment, in 1868 Caucasians. Aspichueta P, editor. PLoS One [Internet]. 2015 Apr 21 [cited 2020 Jan 23];10(4):e0120758. Available from: http://dx.plos.org/10.1371/journal.pone.0120758
19. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol [Internet]. 2015 Apr [cited 2019 Oct 24];44(2):512–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26050253
20. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh P-R, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics [Internet]. Nature Genetics Nature Publishing Group; Jan 28, 2015 p. 1228–35. Available from: http://www.nature.com/articles/ng.3404
21. Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet [Internet]. 2018 May 23 [cited 2018 Oct 19];50(5):693–8. Available from: http://www.nature.com/articles/s41588-018-0099-7
22. Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. Eur J Epidemiol [Internet]. 2018;33(10):947–52. Available from: https://doi.org/10.1007/s10654-018-0424-6

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table

Table 1.
Genetic correlations between smoking traits, type 2 diabetes, and coronary artery disease based on LD Score Regression (20).

Traits		Genetic Correlation (rg)	p-value
Between smoking traits			
Smk Init	Smk Cess	0.391	1.77×10^{-40}
Smk Init	Cig-per-day	0.284	3.96×10^{-17}
Smk Cess	Cig-per-day	0.440	2.29×10^{-42}
Smoking traits vs type 2 diabetes			
Smk Init	type 2 diabetes	0.155	3.13×10^{-19}
Smk Cess	type 2 diabetes	0.196	4.60×10^{-12}
Cig-per-day	type 2 diabetes	0.306	6.96×10^{-23}
Smoking traits vs coronary artery disease (CAD)			
Smk Init	CAD	0.223	1.05×10^{-27}
Smk Cess	CAD	0.187	1.95×10^{-10}
Cig-per-day	CAD	0.294	1.08×10^{-10}

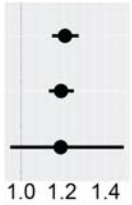
medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure

A

Effect of smoking initiation on type 2 diabetes

# SNPs	Method	OR	95% CI	P-value
341	Inverse variance weighted	1.21	1.15-1.27	1.00×10^{-12}
	Weighted median	1.19	1.13-1.25	6.91×10^{-12}
	MR-Egger	1.19	0.94-1.49	1.34×10^{-1}
	MR-Egger (bias)			8.77×10^{-1}

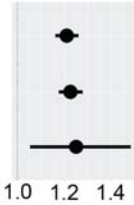


Change in type 2 diabetes odds ratio per 2-fold change in Smoking Initiation

B

Effect of smoking initiation on coronary artery disease

# SNPs	Method	OR	95% CI	P-value
342	Inverse variance weighted	1.21	1.16-1.26	1.97×10^{-20}
	Weighted median	1.23	1.17-1.28	1.41×10^{-20}
	MR-Egger	1.25	1.06-1.48	1.06×10^{-2}
	MR-Egger (bias)			7.00×10^{-2}



Change in CAD odds ratio per 2-fold change in Smoking Initiation

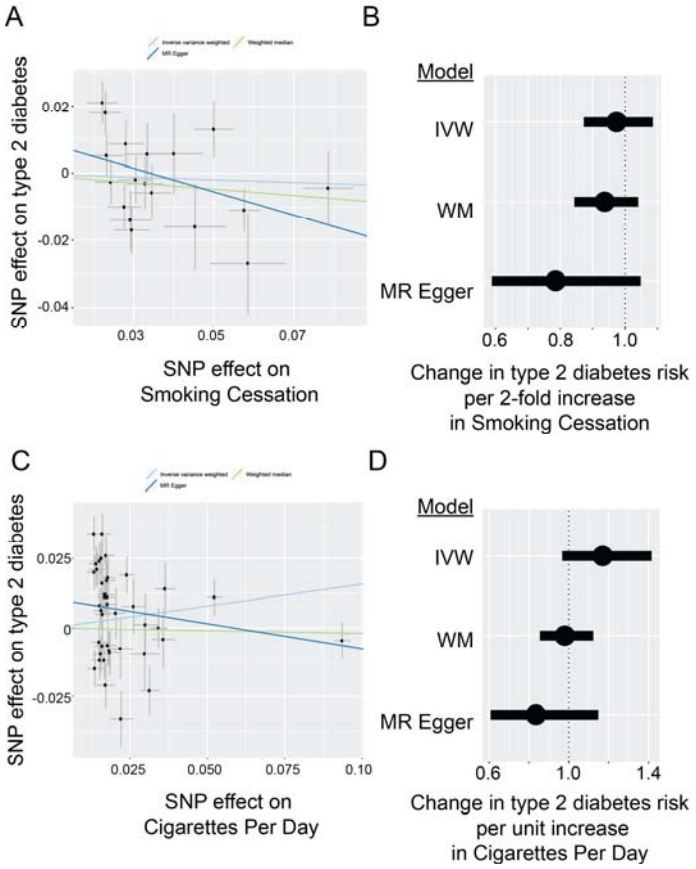
Figure 1.
Two-sample mendelian randomization determines that smoking initiation increases type 2 diabetes risk in similar magnitude to coronary artery disease (CAD) risk.

A. Using a 341-SNP instrumental variable, two-sample MR shows that increased smoking initiation increases type 2 diabetes risk. Odds ratio (OR) estimates and 95% confidence intervals represent changes associated with 2-fold increase in smoking initiation exposure. MR-Egger intercept, a bias measurement, does not deviate significantly from zero. This validates effect (OR) estimates. Forest plot shows change in odds ratio for type 2 diabetes per 2-fold increased smoking initiation risk (mean ± standard error).

B. Using a 342-SNP instrumental variable, two-sample MR shows that increased smoking initiation increases coronary artery disease (CAD) risk. Odds ratio (OR) estimates and 95% confidence intervals represent changes associated with 2-fold increase in smoking initiation exposure. MR-Egger intercept, a bias measurement, does not deviate significantly from zero. This validates effect (OR) estimates. Forest plot shows change in odds ratio for CAD per 2-fold increased smoking initiation risk (mean ± standard error).

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Supplemental Figure



Supplemental Figure 1.
Two-sample mendelian randomization shows no significant effects of smoking cessation or smoking heaviness on type 2 diabetes risk.

A. Using an 18-SNP instrumental variable for smoking cessation, two-sample MR does not define an increased risk of type 2 diabetes. Scatter plot shows effect sizes (mean ± standard error) for smoking cessation and type 2 diabetes for each SNP in the instrumental variable, along with line of best fit for inverse variance weighted, weighted median, and MR Egger regression models.
B. Forest plot showing aggregated effects for smoking cessation on type 2 diabetes, shown as change in type 2 diabetes risk (odds ratio) per 2-fold change in smoking cessation "risk" (mean ± standard error) for inverse variance weighted (IVW), weighted median (WM), and MR Egger regression models.
C. Using a 42-SNP instrumental variable for smoking heaviness (number of cigarettes per day), two-sample MR does not define an increased risk of type 2 diabetes. Scatter plot shows effect sizes (mean ± standard error) for smoking heaviness and type 2 diabetes for each SNP in the instrumental variable, along with line of best fit for inverse variance weighted, weighted median, and MR Egger regression models.
D. Forest plot showing aggregated instrumental variable effects for smoking heaviness on type 2 diabetes, shown as change in type 2 diabetes risk (odds ratio) per unit increase in smoking heaviness (mean ± standard error) for inverse variance weighted (IVW), weighted median (WM), and MR Egger regression models.

medRxiv preprint doi: https://doi.org/10.1101/2020.01.30.20019737; this version posted February 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Supplemental Tables.

Supplementary Table 1.
Instrumental variable data for MR experiments estimating effects of smoking initiation on type 2 diabetes. The rsid (hg19), chromosome, position, effect allele, non-effect allele, effect sizes and standard errors are shown for each SNP.

Supplementary Table 2.
Instrumental variable data for MR experiments estimating effects of smoking cessation on type 2 diabetes. The rsid (hg19), chromosome, position, effect allele, non-effect allele, effect sizes and standard errors are shown for each SNP.

Supplementary Table 3.
Instrumental variable data for MR experiments estimating effects of smoking heaviness (cigarettes per day) on type 2 diabetes. The rsid (hg19), chromosome, position, effect allele, non-effect allele, effect sizes and standard errors are shown for each SNP.

Supplementary Table 4.
Instrumental variable data for MR experiments estimating effects of smoking initiation on coronary artery disease. The rsid (hg19), chromosome, position, effect allele, non-effect allele, effect sizes and standard errors are shown for each SNP.

Supplemental Table 5.
Instrument strength F-statistics and related values from GWAS necessary to perform calculations (http://cnsgenomics.com/shiny/mRnd/, (17)). We calculated genetic risk explained in each exposure using a previously described method Shim *et al* (18).

Supplemental Table 6.
Smoking cessation and smoking heaviness (cigarettes per day, Cig per day) do not reach statistically significant effects on type 2 diabetes risk based on inverse variance weighted (IVW), weighted median (WM) or MR Egger regression metrics. Odds ratio (OR) effects and 95% confidence intervals represent change in type 2 diabetes OR associated with 2-fold increased smoking cessation risk, or change in type 2 diabetes OR per unit increase in smoking heaviness (Cig per day). Significant Egger regression intercept values that deviate significantly from zero (e.g., Cig per day) invalidate effect estimates.