medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 4.0 International license .

A multivariable Mendelian randomisation study exploring the direct effects of nicotine on health compared with the other constituents of tobacco smoke: Implications for e-cigarette use

Jasmine N Khouja (0000-0002-7944-2981), Eleanor Sanderson (0000-0001-5188-5775), Robyn E Wootton (0000-0003-3961-3202), Amy E Taylor, Marcus R Munafò (0000-0002-4049-993X)

School of Psychological Science, 12a Priory Road, University of Bristol, Bristol, United Kingdom, BS8 1TU, Jasmine N Khouja, Lecturer. Bristol Medical School: Population Health Sciences, Oakfield House, University of Bristol, Bristol, United Kingdom, BS8 2BN, Eleanor Sanderson, Senior Research Associate. Bristol Medical School: Population Health Sciences, Oakfield House, University of Bristol, Bristol, United Kingdom, BS8 2BN, Amy E Taylor, Research Fellow. Nic Waals Institute, Lovisenberg Diaconal Hospital, Ullevålsveien 2, 0165, Oslo, Norway, Robyn E Wootton, Postdoctoral Fellow. NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and School of Psychological Science, 12a Priory Road, University of Bristol, Bristol, United Kingdom, BS8 1TU, Marcus R Munafò, Professor of Biological Psychology.

Correspondence to: Jasmine N. Khouja; jasmine.khouja@bristol.ac.uk; 12a Priory Road, University of Bristol, Bristol, UK, BS8 1TU

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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/10.1101/2021.01.12.21249493; this version posted January 15, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 4.0 International license .

Abstract

Objectives Given the popularity of e-cigarettes, and the lack of longitudinal evidence regarding their safety, novel methods are required to explore potential health effects resulting directly from nicotine use. The aim of this study was to explore the direct effects of nicotine compared with the other constituents of tobacco smoke on health outcomes associated with smoking.

Design Observational study, using Mendelian randomisation and multivariable Mendelian randomisation analyses of summary data.

Setting Summary data from two previous genome-wide association studies, and summary data generated from UK Biobank, a prospective cohort study.

Participants N = 337,010 individuals enrolled in UK Biobank, and a total of N = 341,882 individuals from two previous genome-wide association studies**.**

Main outcome measures We explored the effect of cotinine levels (as a proxy for nicotine exposure) and smoking heaviness (to capture cigarette smoke exposure) on body mass index (BMI), chronic obstructive pulmonary disease (COPD), forced vital capacity (FVC), forced expiratory volume (FEV-1), coronary heart disease (CHD), and heart rate.

Results In multivariable Mendelian randomisation analyses, there was weak evidence to suggest that increased cotinine levels may cause increased heart rate among current smokers (β = 0.50 bpm, 95% CI -0.06 to 1.05). There was stronger evidence to suggest that increased smoking heaviness causes decreased BMI among current smokers (β = -1.81 kg/m^2, 95% CI -2.64 to -0.98), as well as increased risk of COPD, decreased FEV-1 and FVC, and increased heart rate among ever and current smokers. We also found evidence to suggest that increased smoking heaviness causes increased risk of CHD among ever smokers.

Conclusions Our combined findings are consistent with smoking-related health outcomes being caused by exposure to the non-nicotine components of tobacco smoke.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

Keywords: smoking; nicotine; health; mendelian randomisation

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

Introduction

Of an estimated 3.6 million e-cigarette users in Great Britain, 22% use e-cigarettes to help them stop smoking.[1] Although current evidence suggests that e-cigarettes may reduce harm by aiding smoking cessation,[2-4] the long-term health effects of nicotine exposure via e-cigarette use remain unknown. In contrast, the long-term health outcomes of smoking are well-known, given the abundance of observational evidence demonstrating associations between smoking and health issues such as chronic obstructive pulmonary disease (COPD), coronary heart disease (CHD) and poor lung function.[5-8] Consistent evidence across many observational studies provides strong support for a causal effect,[9] which is further supported by genome-wide association studies that identify smoking-related genetic variants when examining these outcomes.[10-13] However, it remains unclear which constituents of tobacco smoke (e.g., nicotine, carbon monoxide) negatively impact health, or have the largest effects.

Until e-cigarettes became widely available in 2007, nicotine replacement therapy (NRT) was the primary source of nicotine without tobacco. However, long-term NRT use is rare among ex-smokers [14] and non-smokers;[15] consequently, there is little evidence on the long-term effects of nicotine use when not consumed in tobacco products. Given that a randomised controlled trial of long-term nicotine use would be unethical, we require alternative methods to estimate causal consequences of nicotine use. Mendelian randomisation (MR) is a method which is often used to infer causality, particularly where a randomised controlled trial would be unethical or impossible.[16] The method assumes that the laws of Mendelian genetics (segregation and independent assortment) are held at a population level i.e., a random assortment of genes are transferred from parents to their offspring.[17] For example, if an individual inherits the rs16969968 genetic variant which predisposes them to be more tolerant of nicotine, then they are likely to be a heavier smoker (i.e., smoke one more cigarette per day per risk allele) on average than an individual who did not inherit those variants.[18] The inheritance of these genetic variants is mostly independent of confounding factors

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

which often distort observational evidence, and therefore mimics the randomisation process in a randomised controlled trial, reducing issues of both confounding and reverse causality.[17 19 20] The MR method estimates the total causal effect of one exposure on one outcome. For example, to explore the potential harm of using nicotine-containing products (e.g., e-cigarettes), we could use MR methods to estimate the total effect of e-cigarette use on COPD. However, to conduct MR analysis, we require large genome-wide association studies (GWAS) of the exposure, to identify genetic variants that can be used as proxies for this exposure. At present, there are no published large GWAS of e-cigarette use, and no consortia with sufficient numbers of e-cigarette users to support GWAS analysis. Furthermore, e-cigarette use and smoking are highly correlated[1] and may share a genetic aetiology,[21 22] so to ensure any associations found are not due to confounding effects of smoking, the GWAS should be restricted to never-smokers, but few never-smokers regularly vape.[1 23 24]

Multivariable MR (MVMR) is an extension of the MR method; rather than calculating the total effect, MVMR is used to explore the direct causal effect of two or more exposures on an outcome.[25 26] When two exposures are related, MVMR can estimate the effect of one exposure on an outcome while accounting for the effect of the other exposure on the outcome (i.e., the direct effect) even when there is overlap in the genetic effects on the two exposures. As cigarettes contain nicotine, smoke exposure and nicotine intake are highly correlated; therefore, MVMR is a suitable method to explore the direct effects of nicotine versus the direct effect of the other constituents of tobacco smoke on smoking-related health outcomes. GWAS have previously identified genetic variants associated with smoking heaviness as well as cotinine – a highly-specific biomarker which captures recent nicotine exposure given that 70-80% of nicotine is rapidly metabolised into cotinine. [27] By using these genetic variants as proxies for nicotine (G_N) and smoking (G_S) – including the genetic variants that predict both (G_{SN}) – in an MVMR analysis (Figure 1a), we can explore the direct effects of cotinine while taking into account the effects of smoke exposure (Figure 1b) and vice versa (Figure 2c). The total effects of smoking heaviness on health outcomes include the effects of cotinine on

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

health outcomes, but by exploring the direct effects of smoking heaviness while controlling for the direct effects of cotinine, we can observe the effects of the remaining constituents of tobacco smoke (Figure 1c). In other words, among smokers, we can identify the health effects caused by nicotine versus the health effects caused by the other constituents of tobacco smoke exposure. The aim of this study was therefore to employ MVMR methods to explore the direct effects of nicotine compared with the other constituents of tobacco smoke on health outcomes known to be caused by smoking.

Methods

Data Sources

The data sources for the exposures (cotinine and smoking heaviness) and health outcomes are shown in Figure 2. The data obtained from these sources are described as either summary-level or individual-level. Summary-level data contain only the overall genetic association with the exposure and outcome for the whole sample and can be used to identify suitable genetic instruments and the effect sizes of the instrument-phenotype association for inclusion in MR and MVMR analysis. Individual-level data consist of genetic, exposure and outcome data for all individual participants with which genetic associations can be calculated. Where individual-level data are provided, summary-level data can be generated for further analysis (Figure 2).

The Cotinine Consortium. Ware, et al. [28] report summary-level statistics from a GWAS meta-analysis of cotinine levels (per standard deviation change) among daily smokers of European Ancestry (data available at: https://doi.org/10.5523/bris.182rhz19hg3lz1172a7yfcap9v). SNPs were reported as independent if they reached genome-wide significance using an iterative process of conditional analyses. Further information about this GWAS can be found in the supplementary material (Supplementary Note 1).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

GSCAN. The GWAS & Sequencing Consortium of Alcohol and Nicotine use (GSCAN) reports summary-level statistics from a GWAS of smoking heaviness[29] (data available at: https://doi.org/10.13020/3b1n-ff32). SNPs were reported as independent if they explain additional variance in conditional analyses using a partial correlation-based score statistic.[30] Further information about this GWAS can be found in the supplementary material (Supplementary Note 2).

UK Biobank. We obtained individual-level data from UK Biobank, a population-based health research resource consisting of approximately 500,000 people, aged between 38 years and 73 years, who were recruited between the years 2006 and 2010 from across the UK.[31] With a particular focus on identifying determinants of human diseases in middle-aged and older individuals, participants provided a wide range of health information (data available at www.ukbiobank.ac.uk). A full description of the study design, participants and quality control (QC) methods have been described in detail previously.[31 32] UK Biobank received ethics approval from the Research Ethics Committee (REC reference for UK Biobank is 11/NW/0382). Written informed consent was obtained from participants prior to their participation in UK Biobank. After restricting the sample to individuals of White British ancestry[33] and excluding those with mismatched sex, with missing array data, who were related or withdrew their consent to participate, the sample size was 337,010.[32]

Health Outcomes

Body mass index (BMI) and heart rate (beats per minute) were measured during a UK Biobank Assessment Centre visit. We identified COPD cases as participants who self-reported a doctor's diagnosis of COPD. Forced expiratory volume in 1 second (FEV-1) and forced vital capacity (FVC) were measured using a Vitalograph spirometer. CHD diagnosis was determined using linked hospital admission data (ICD codes relating to Ischemic Heart Disease). Further information regarding each health outcome (including UK Biobank field IDs) can be found in the supplementary material (Supplementary Note 3).

Smoking Status

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

In UK Biobank, smoking status was categorised as never, previous and current smoking (field ID 20116). From this variable, we derived an 'ever smokers' variable which was defined as currently or having previously smoked occasionally, most days or daily (i.e., having smoked more than just once or twice). Current smoking was defined as currently smoking occasionally, most days or daily. Former smoking was defined as not currently smoking but having previously smoked occasionally, most days or daily (i.e., more than just once or twice). Those who have tried smoking once or twice or who have never smoked were categorised as never smokers.

Generated Summary Statistics

Using individual-level data from UK Biobank, we generated summary-level data by regressing each SNP on each of the health outcomes, adjusting for 10 principal components of population stratification. As shown in Figure 2, four datasets were generated according to smoking status: ever smokers (including current and former smokers; n = 151,809), current smokers (n = 33,354), former smokers (n = 118,455), and never smokers (n = 184,016).

Statistical Analysis

Analyses were carried out in Stata 15.1.[34]

Selection of genetic variants. Genetic variants related to the phenotype of interest (cotinine levels or smoking heaviness) were selected for inclusion in the analysis based on the reported results of the relevant GWAS (Figure 2). SNPs that were independent of any other SNP associations at the genome-wide significant level ($p < 5x10^{-8}$) were selected for inclusion – 55 SNPs were identified as associated with smoking heaviness[29] and 3 SNPs were identified as associated with cotinine levels.[28]

For MVMR analyses, all included SNPs (i.e., those relating to smoking heaviness as well as those relating to cotinine levels) must also be independent of each other, so an additional clumping stage was added to ensure overall SNP independence (LD $R^2 < 0.1$, clumping window > 500 kb).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

Due to the limited number of independent SNPs associated with cotinine levels at the genome-wide significant threshold (n = 2), the significance threshold used for inclusion of cotinine SNPs was lowered to $p < 5 \times 10^{-6}$ for cotinine SNPs included in the main analyses. Where a SNP was identified for inclusion but was not available in either of the other summary data sets (i.e., available in the Cotinine Consortium summary data but not available in the GSCAN summary data or UK Biobank data), we selected proxy SNPs with a minimum linkage disequilibrium (LD) R^2 of 0.8. Details of the SNPs included in each analysis, and proxies used, are provided in Supplementary Table 1 and Supplementary Note 4.

Instrument strength and validity was tested using the conditional F-statistic for MVMR and Cochran *Q* statistic.[35 36] As a general rule, the F-statistic should be greater than 10 and *Q* estimates should be less than the number of SNPs included in the model.

Multivariable Mendelian randomisation. We explored the direct effects of cotinine levels and smoking heaviness on each health outcome (BMI, COPD, FEV-1, FVC, CHD and heart rate) individually using MVMR. To conduct the MVMR, we used summary data from the Cotinine Consortium and GSCAN, and summary data generated using individual-level data from UK Biobank (binary outcomes were estimated using logistic regressions, and continuous outcomes using linear regressions). We repeated these analyses using two complimentary methods – MVMR-IVW and MVMR-Egger.[26] All of the SNPs included in these analyses were associated with either cotinine levels or smoking heaviness (or both).

To explore the recoverable and long-term outcomes of smoking, this analysis was restricted to (1) ever smokers, and (2) current smokers. In supplementary analyses, we additionally stratified the analysis by former smoking status to further explore recoverable effects, and we stratified the analysis by ever smoking status to explore potential horizontal pleiotropy – effects observed among never smokers could indicate horizontally pleiotropic effects (i.e., the included SNPs influencing the outcome directly, or via another phenotype, but not through smoking), misreporting of smoking

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

status, or population stratification. Horizontally pleiotropic genetic variants are not valid instruments in MR analyses.

Univariable Mendelian randomisation. For comparison with the main analysis, we considered the total effect of both cotinine levels and smoking heaviness on each health outcome using MR. Details of the univariable MR analysis methods can be found in Supplementary Note 5.

Public and Patient Involvement

Given the nature of the study (i.e., the use of secondary and summary data), there was no input from patients or the public in the design or implementation of this study.

Results

Descriptive Statistics

Of the 337,010 individuals with available data in UK Biobank, 54% were male, the average age was 57 years, and the average BMI was 27.39. A total of 1,245 (1%) had a diagnosis of COPD and 28,652 (9%) had a diagnosis of CHD. Average FEV-1 was 2.87 litres, average FVC was 3.80 litres, and average heart rate was 68.98 bpm. A total of 184,016 (55%) were never smokers and 151,809 (45%) were ever smokers, of whom 33,354 (10%) were current smokers and 118,455 (35%) were former smokers. A total of 1,185 UK Biobank participants who preferred not to state their smoking status were excluded.

Multivariable Mendelian Randomisation

The complete results of the MVMR-IVW analysis exploring the direct effects of cotinine and cigarettes per day (n = 54 SNPs) on health outcomes are displayed in Supplementary Table 2. Results are presented per standard deviation (SD) increase in the exposure phenotype (i.e., cotinine/cigarettes per day).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

The Cochran's Q statistics were greater than the number of SNPs included (N = 54) in the majority of the models, indicating heterogeneity (Supplementary Table 2). Therefore, we also present the main analysis using a pleiotropy robust method, MVMR-Egger (Supplementary Table 3), which gives estimates that are robust to directional pleiotropy under the assumption that this pleiotropy is uncorrelated with the strength of association between the SNP and the exposure.[37] However, a limitation of this approach is limited statistical power compared to MVMR-IVW.

Instrument Strength. Instrument strength was calculated using the two-sample conditional F-statistic.[36] The conditional F-statistic for MVMR indicates instrument strength of each exposure when accounting for the prediction of other exposures in the model (i.e., whether the SNPs jointly predict smoking heaviness after predicting cotinine levels).[36] This indicated that the SNPs included in the analysis are strong instruments for assessing the direct effects of smoking heaviness while accounting for the effect of cotinine levels (F = 21.66). However, this also indicated that the SNPs may not be strongly associated with cotinine levels while accounting for the effect of smoking heaviness (F = 6.83). These F-statistics were calculated with the use of a less stringent threshold ($p < 5 \times 10^{-6}$) for the inclusion of SNPs associated with cotinine levels. Use of the less stringent threshold improved the instrument strength compared with the genome-wide significant threshold by adding more independent SNPs that are only associated with cotinine and not with smoking heaviness ($p < 5 \times 10^{-8}$, F for smoking heaviness = 17.53; F for cotinine = 3.36), which supports the use of the less stringent threshold for the main analysis.

Direct effects of cotinine levels on health outcomes. When taking into account the effects of smoking heaviness in the MVMR-IVW analysis, there was no clear evidence of an effect of cotinine on heart rate among ever smokers (Supplementary Table 2). However, there was some weak evidence to suggest that increased cotinine levels cause increased heart rate among current smokers (β = 0.50 bpm, 95% CI -0.06 to 1.05 per SD increase in cotinine levels), indicating some evidence of an acute, recoverable effect. There was no clear evidence of any other effect of cotinine levels on

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

smoking-related health outcomes (Supplementary Table 2). The results were similar in the MVMR-Egger analysis, but there was weak evidence to suggest that cotinine lowers the risk of CHD among current smokers (OR = 0.86, 95% CI -0.74 to 1.01 per SD increase in cotinine levels).

Direct effects of smoking heaviness on health outcomes. When taking into account the effect of cotinine levels in the MVMR-IVW analysis, there was no clear evidence to suggest an effect of increased smoking heaviness on BMI among ever smokers (Supplementary Table 2), but there was evidence to suggest that increased smoking heaviness decreases BMI among current smokers (β = -1.81 kg/m^2, 95% CI -2.64 to -0.98 per SD increase in cigarettes smoked per day). In the MVMR-Egger analysis, there was evidence to suggest that increased smoking heaviness decreases BMI among ever smokers (β = -1.02 kg/m^2, 95% CI -2.01 to -0.03 per SD increase in cigarettes smoked per day). The results of the MVMR-Egger analysis were similar for current smokers.

When taking into account the effect of cotinine levels in the MVMR-IVW analysis, there was evidence to suggest that increased smoking heaviness causes increased risk of COPD among ever smokers (OR = 7.32, 95% CI 3.60 to 14.88 per SD increase in cigarettes smoked per day) and current smokers (OR = 29.37, 95% CI 9.68 to 89.12 per SD increase in cigarettes smoked per day). The results of the MVMR-Egger were similar for ever and current smokers (Supplementary Table 3).

There was also evidence to suggest that increased smoking heaviness causes decreased FEV-1 and FVC among ever smokers (β = -0.22 litres, 95% CI -0.29 to -0.15; β = -0.19 litres, 95% CI -0.28 to -0.09 per SD increase in cigarettes smoked per day respectively) and current smokers (β = -0.34 litres, 95% CI -0.48, -0.20; β = -0.24 litres, 95% CI -0.41 to -0.06 per SD increase in cigarettes smoked per day respectively) in the MVMR-IVW analysis. However, there was no clear evidence of an effect of smoking heaviness on FEV-1 or FVC in the MVMR-Egger analysis except for some weak evidence of an effect on FEV-1 among ever smokers (β = -0.11 litres, 95% CI -0.24 to 0.01 per SD increase in cigarettes smoked per day; Supplementary Table 3).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

There was evidence to suggest that increased smoking heaviness causes increased risk of CHD among ever smokers (OR = 1.36, 95% CI 1.05 to 1.79 per SD increase in cigarettes smoked per day), but not among current smokers (Supplementary Table 2) in the MVMR-IVW analysis. There was no clear evidence of an effect of smoking heaviness among ever smokers in the MVMR-Egger analysis (Supplementary Table 3).

There was evidence to suggest that increased smoking heaviness raises heart rate among ever smokers (β = 1.83, 95% CI 0.88 to 2.79 per SD increase in cigarettes smoked per day) and current smokers (β = 3.00, 95% CI 1.41 to 4.59 per SD increase in cigarettes smoked per day) in the MVMR-IVW analysis. In the MVMR-Egger analysis however, there was no clear evidence of an effect of smoking heaviness on heart rate (Supplementary Table 3).

Comparing the total and direct effects. The results of the univariable analyses (i.e., the total effects) are reported in Supplementary Note 6 and in Supplementary Tables 4 and 5. Among ever smokers, the differences between the total and direct effects of smoking heaviness were negligible for the IVW (Figure 3) and Egger (Supplementary Figure 1) analyses. Among current smokers, the differences between the total and direct effects of smoking heaviness were also negligible for the IVW (Figure 4) and Egger (Supplementary Figure 2) analyses. The effect estimates were similar in magnitude and direction, and the confidence intervals overlapped.

Sensitivity and supplementary analysis. The results did not substantially differ from the reported results when the more stringent threshold ($p < 5 \times 10^{-8}$) was used for the IVW method in the multivariable (Supplementary Table 6) or univariable analyses (Supplementary Table 7). Tests of the weighted regression dilution, instrument validity, heterogeneity and directional pleiotropy for the univariable analyses can be found in Supplementary Tables 8-10. All the effects described in the above IVW analysis also had consistent evidence from the sensitivity analyses. The results for former smokers and never smokers are shown in Supplementary Tables 2 and 3. There was some evidence to suggest pleiotropic effects (i.e., there was some evidence of an effect of genetic propensity to

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

heavier smoking on some health outcomes). The effects found are reported in full in Supplementary Notes 7 and 8.

Discussion

Our results confirm the known effects of smoking on health. Critically, the direct effects of smoking heaviness are similar to the total effects of smoking heaviness, suggesting that these health outcomes are not caused by nicotine per se, but by the other non-nicotine constituents of cigarette smoke. In contrast, there is little clear evidence of a direct effect of nicotine on smoking-related health outcomes, although this could be due to a lack of statistical power. Combined, this evidence indicates that nicotine is likely to have relatively little impact on these health outcomes, certainly compared with the impact of the other constituents of tobacco smoke, which appear to cause numerous negative health effects related to smoking.

When interpreting these results, it is important to consider the validity of the instruments used. The conditional F-statistics indicated that the instrument used as a proxy for smoking heaviness was strong, but the instrument used as a proxy for nicotine was weak.[36] In univariable MR, the F-statistic simply indicates the instrument strength of the single exposure; however, the conditional F-statistic in the MVMR context indicates instrument strength of each exposure when accounting for the prediction of other exposures in the model (i.e., whether the SNPs jointly predict smoking heaviness after predicting cotinine levels). Therefore, the conditional F-statistic indicates that we can be confident in the estimate of the direct effect of smoking heaviness when taking into account nicotine exposure (i.e., the effect of other constituents of tobacco smoke aside from nicotine). However, we cannot be as confident in the estimate of the direct effect of nicotine on smoking-related health outcomes. The main genetic variant identified in the cotinine GWAS (rs10851907) is in LD with rs16969968 (a known functional variant associated with smoking heaviness) which could explain why the instrument is conditionally weak.[28]

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

For the most part, the differences between the total and direct effects of smoking heaviness were small, implying that nicotine has little direct impact on smoking-related health outcomes. The results are somewhat in line with previous evidence which suggests that nicotine may have an effect on resting heart rate.[38] However, the results do not suggest that nicotine use without exposure to tobacco smoke has a direct effect on CHD, BMI, lung function, or COPD. In contrast, there was evidence of a direct effect of the other constituents of tobacco smoke on the selected health outcomes which have previously been shown to be associated with smoking.[5-8] Interestingly, we also found some evidence of an effect of increased genetic propensity to smoke and nicotine on BMI, COPD and heart rate among never smokers which is indicative of pleiotropic effects.

Strengths and Weaknesses of the Study

This study is the first to explore the long-term effects of nicotine use among smokers while considering the direct effects of other constituents of tobacco smoke (and vice versa). We have employed a novel method (MVMR) to explore the causal effect of nicotine on potential health outcomes in order to give an indication of possible future health consequences of long-term nicotine-containing e-cigarette use. However, this study is not without limitations. First, there are issues interpreting findings where the number of cigarettes per day are used as a proxy for smoke exposure. As described by Taylor and colleagues,[39] the number of cigarettes smoked per day is often used to determine lifetime smoke exposure, but there are individual differences in smoking topography (i.e., number of puffs taken per cigarette, average volume per puff etc.) which are not captured by measures of cigarettes per day, meaning measures of cigarettes per day may not adequately capture smoke exposure. Second, pleiotropy may have impacted these results. Interestingly, we found evidence of effects of genetic propensity to heavier smoking on BMI and COPD among those who have never smoked before in the MVMR-IVW analysis (Supplementary Table 2) and we also found some evidence of an effect of genetic propensity to use nicotine on heart rate among never smokers in the MVMR-Egger analysis (Supplementary Table 3). The effect

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

estimates among never smokers cannot be meaningfully interpreted as we know that never smokers do not smoke any cigarettes per day despite being predisposed to heavier smoking. Therefore, evidence for an effect in never smokers (along with a high Cochran's Q statistic) is indicative of horizontal pleiotropy (i.e., the genetic variants influencing smoking heaviness also separately influence BMI/COPD through a pathway other than smoking). Additionally, the MR-Egger test of directional pleiotropy indicated directional pleiotropic effects in the relationship between smoking heaviness and health outcomes, particularly among ever smokers. As there is a smaller sample of current smokers (but the size of the intercept is similar to ever smokers), there may be some pleiotropic effects for current smokers which have not been detected due to a lack of statistical power. Third, cotinine is not a perfect biomarker of nicotine; nicotine metabolism (and therefore cotinine levels) can be affected by a person's age, gender, and even diet.[27] Additionally, BMI (one of our health outcomes of interest) can impact metabolism in general[40] so could influence cotinine metabolism. Consequently, there is likely to be some measurement error in the estimates used to determine which SNPs are independently associated with nicotine consumption. Although cotinine is not a perfect biomarker of nicotine, it is unlikely that a GWAS of nicotine will become available given its short half-life (~2 hours) which makes direct measurement difficult.[27] Fourth, the conditional F-statistic indicated that the estimates of the direct effects of nicotine exposure on the health outcomes are likely to suffer from weak instrument bias and should therefore be interpreted with caution. The GWAS of cotinine[28] was based on a relatively small sample size (N = 4,548) compared to the GWAS of smoking heaviness (N = 120,744)[29] and may have lacked power to detect some influential SNPs.

Future Research and Implications

As more GWAS summary data become available (e.g., Buchwald and colleagues),[41] a larger scale GWAS of cotinine may reveal more independent SNPs which can be used as an instrument for nicotine exposure. Future research could extend on these findings using a stronger instrument for

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

cotinine (if available) which would allow for clearer interpretation of the causal effect of long-term nicotine use. The current evidence suggests that consuming nicotine without tobacco smoke (e.g., via e-cigarettes rather than cigarettes) may reduce the risk of developing smoking-related diseases. If the findings are supported by studies using a stronger instrument as a proxy for cotinine, then policies which encourage smokers to switch to e-cigarettes could lessen the health burden of smoking-related disease on public health care systems such as the NHS.

Conclusion

In conclusion, although we found clear evidence of a direct causal effect of exposure to the other constituents of cigarette smoke aside from cotinine on a range of health outcomes, we only observed evidence of a direct effect of cotinine on heart rate. Despite potential weak instrument bias in the estimates of the direct effect of cotinine, we can cautiously infer that nicotine use via cigarettes has little impact on the selected smoking-related health outcomes because there is little difference between the total effects of smoking heaviness (when nicotine exposure is not taken into account) and the direct effects of smoking heaviness (when the direct effect of nicotine is taken into account). Although we did not identify any strong effects of nicotine on health in this study, nicotine may still have a small influence on health independent of smoking. However, nicotine does not appear to be the main cause of the negative effects of cigarettes on these specific health outcomes. This suggests that long-term use of nicotine without the other constituents of cigarette smoke (e.g., vaping or NRT use) would result in fewer of the selected negative health outcomes than long-term smoking. However, the impact of nicotine requires further investigation and future studies should further explore the role of nicotine with a stronger instrument if relevant data become available.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

Funding

This work was supported by the Medical Research Centre Integrative Epidemiology Unit at the University of Bristol [grant number MC_UU_00011/1, MC_UU_00011/7]. This research was also supported by the NIHR Bristol Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care. REW is supported by a postdoctoral fellowship from the South-Eastern Regional Health Authority (2020024). This publication is the work of the authors and JNK, ES, REW, AET and MRM who had full access to all of the data (including statistical reports and tables) and will serve as guarantors for the contents of this paper. The authors take responsibility for the integrity of the data and the accuracy of the data analysis. The funder played no role in the design of the study, the data collection, data analysis, interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

Conflicts of interest

There are no conflicts of interest to declare.

Transparency declaration

The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

References

1. Action on Smoking and Health. Use of e-cigarettes (vapourisers) among adults in Great Britain, 2019.
2. Hajek P, Phillips-Waller A, Przulj D, et al. A Randomized Trial of E-Cigarettes versus Nicotine-Replacement Therapy. *New Engl J Med* 2019;380(7):629-37. doi: 10.1056/NEJMoa1808779
3. Hartmann-Boyce J, McRobbie H, Bullen C, et al. Electronic cigarettes for smoking cessation. *Cochrane Database of Systematic Reviews* 2016(9) doi: 10.1002/14651858.CD010216.pub3
4. McNeill A, Brose LS, Calder R, et al. E-cigarettes: an evidence update. In: ENgland PH, ed., 2015.
5. West R. Tobacco smoking: Health impact, prevalence, correlates and interventions. *Psychol Health* 2017;32(8):1018-36. doi: 10.1080/08870446.2017.1325890 [published Online First: 2017/05/30]
6. Jayes L, Haslam PL, Gratziou CG, et al. SmokeHaz: Systematic Reviews and Meta-analyses of the Effects of Smoking on Respiratory Health. *Chest* 2016;150(1):164-79. doi: 10.1016/j.chest.2016.03.060 [published Online First: 2016/04/23]
7. Das SK. Harmful health effects of cigarette smoking. *Mol Cell Biochem* 2003;253(1-2):159-65. doi: 10.1023/a:1026024829294 [published Online First: 2003/11/19]
8. Khan A, Shabbir K, Ansari JK, et al. Comparison of forced expiratory volume in one second (FEV1) among asymptomatic smokers and non-smokers. *J Pak Med Assoc* 2010;60(3):209-13. [published Online First: 2010/03/17]
9. Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med* 1965;58:295-300. [published Online First: 1965/05/01]
10. Zhang R, Chu M, Zhao Y, et al. A genome-wide gene-environment interaction analysis for tobacco smoke and lung cancer susceptibility. *Carcinogenesis* 2014;35(7):1528-35. doi: 10.1093/carcin/bgu076 [published Online First: 2014/03/25]
11. Hobbs BD, de Jong K, Lamontagne M, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet* 2017;49(3):426-32. doi: 10.1038/ng.3752 [published Online First: 2017/02/07]
12. Polfus LM, Smith JA, Shimmin LC, et al. Genome-wide association study of gene by smoking interactions in coronary artery calcification. *PLoS One* 2013;8(10):e74642. doi: 10.1371/journal.pone.0074642 [published Online First: 2013/10/08]
13. Gage SH, Davey Smith G, Ware JJ, et al. G = E: What GWAS Can Tell Us about the Environment. *PLoS Genet* 2016;12(2):e1005765. doi: 10.1371/journal.pgen.1005765 [published Online First: 2016/02/13]
14. Shahab L, Dobbie F, Hiscock R, et al. Prevalence and Impact of Long-term Use of Nicotine Replacement Therapy in UK Stop-Smoking Services: Findings From the ELONS Study. *Nicotine Tob Res* 2017;20(1):81-88. doi: 10.1093/ntr/ntw258 [published Online First: 2016/09/25]
15. Jackson SE, Hill E, Shahab L, et al. Prevalence and correlates of long-term e-cigarette and nicotine replacement therapy use: a prospective study in England. *BMJ Open* 2019;9(10):e029252. doi: 10.1136/bmjopen-2019-029252 [published Online First: 2019/10/13]
16. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *International Journal of Epidemiology* 2003;32(1):1-22. doi: 10.1093/ije/dyg070
17. Nitsch D, Molokhia M, Smeeth L, et al. Limits to causal inference based on Mendelian randomization: a comparison with randomized controlled trials. *Am J Epidemiol* 2006;163(5):397-403. doi: 10.1093/aje/kwj062 [published Online First: 2006/01/18]
18. Ware JJ, van den Bree MBM, Munafò MR. Association of the CHRNA5-A3-B4 Gene Cluster With Heaviness of Smoking: A Meta-Analysis. *Nicotine & Tobacco Research* 2011;13(12):1167-75. doi: 10.1093/ntr/ntr118
19. Swanson SA, Tiemeier H, Ikram MA, et al. Nature as a Trialist?: Deconstructing the Analogy Between Mendelian Randomization and Randomized Trials. *Epidemiology* 2017;28(5):653-59. doi: 10.1097/ede.0000000000000699 [published Online First: 2017/06/08]

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

20. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;362:k601. doi: 10.1136/bmj.k601
21. Allegrini AG, Verweij KJH, Abdellaoui A, et al. Genetic Vulnerability for Smoking and Cannabis Use: Associations With E-Cigarette and Water Pipe Use. *Nicotine Tob Res* 2019;21(6):723-30. doi: 10.1093/ntr/nty150 [published Online First: 2018/07/28]
22. Khouja JN, Wootton RE, Taylor AE, et al. Association of genetic liability to smoking initiation with e-cigarette use in young adults. *medRxiv* 2020:2020.06.10.20127464. doi: 10.1101/2020.06.10.20127464
23. Bauld L, MacKintosh AM, Eastwood B, et al. Young People's Use of E-Cigarettes across the United Kingdom: Findings from Five Surveys 2015-2017. *Int J Environ Res Public Health* 2017;14(9) doi: 10.3390/ijerph14090973 [published Online First: 2017/08/30]
24. Hammond D, Reid JL, Rynard VL, et al. Prevalence of vaping and smoking among adolescents in Canada, England, and the United States: repeat national cross sectional surveys. *BMJ* 2019;365:l2219. doi: 10.1136/bmj.l2219 [published Online First: 2019/06/22]
25. Lawlor DA, Wade K, Borges MC, et al. A Mendelian Randomization dictionary: Useful definitions and descriptions for undertaking, understanding and interpreting Mendelian Randomization studies. 2019 doi: https://doi.org/10.31219/osf.io/6yzs7
26. Sanderson E, Davey Smith G, Windmeijer F, et al. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol* 2019;48(3):713-27. doi: 10.1093/ije/dyy262 [published Online First: 2018/12/12]
27. Benowitz NL, Hukkanen J, Jacob P, 3rd. Nicotine chemistry, metabolism, kinetics and biomarkers. *Handb Exp Pharmacol* 2009(192):29-60. doi: 10.1007/978-3-540-69248-5_2 [published Online First: 2009/02/03]
28. Ware JJ, Chen X, Vink J, et al. Genome-Wide Meta-Analysis of Cotinine Levels in Cigarette Smokers Identifies Locus at 4q13.2. *Sci Rep* 2016;6:20092. doi: 10.1038/srep20092 [published Online First: 2016/02/03]
29. Liu MZ, 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(2):237-+. doi: 10.1038/s41588-018-0307-5
30. Jiang Y, Chen S, McGuire D, et al. Proper conditional analysis in the presence of missing data: Application to large scale meta-analysis of tobacco use phenotypes. *PLOS Genetics* 2018;14(7):e1007452. doi: 10.1371/journal.pgen.1007452
31. Allen NE, Sudlow C, Peakman T, et al. UK biobank data: come and get it. *Sci Transl Med* 2014;6(224):224ed4. doi: 10.1126/scitranslmed.3008601 [published Online First: 2014/02/21]
32. Mitchell R, Hemani G, Dudding T, et al. UK Biobank Genetic Data: MRC-IEU Quality Control, version 2. 2019 doi: https://doi.org/10.5523/bris.1ovaau5sxunp2cv8rcy88688v
33. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562(7726):203-09. doi: 10.1038/s41586-018-0579-z [published Online First: 2018/10/12]
34. Stata Statistical Software [program]. 15.1 version. College Station, TX: StataCorp LLC, 2017.
35. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: A review. *Res Synth Methods* 2019;10(4):486-96. doi: 10.1002/jrsm.1346 [published Online First: 2019/03/13]
36. Sanderson E, Spiller W, Bowden J. Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomisation. *bioRxiv* 2020:2020.04.02.021980. doi: 10.1101/2020.04.02.021980
37. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Statistics in Medicine* 2017;36(29):4705-18. doi: 10.1002/sim.7492

medRxiv preprint doi: https://doi.org/10.1101/2021.01.12.21249493; this version posted January 15, 2021. 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 4.0 International license .

38. Linneberg A, Jacobsen RK, Skaaby T, et al. Effect of Smoking on Blood Pressure and Resting Heart Rate. *Circulation: Cardiovascular Genetics* 2015;8(6):832-41. doi: doi:10.1161/CIRCGENETICS.115.001225
39. Taylor AE, Davies NM, Ware JJ, et al. Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. *Econ Hum Biol* 2014;13:99-106. doi: 10.1016/j.ehb.2013.12.002 [published Online First: 2014/01/07]
40. Wurtz P, Wang Q, Kangas AJ, et al. Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. *PLoS Med* 2014;11(12):e1001765. doi: 10.1371/journal.pmed.1001765 [published Online First: 2014/12/10]
41. Buchwald J, Chenoweth MJ, Palviainen T, et al. Genome-wide association meta-analysis of nicotine metabolism and cigarette consumption measures in smokers of European descent. *Mol Psychiatry* 2020 doi: 10.1038/s41380-020-0702-z [published Online First: 2020/03/12]

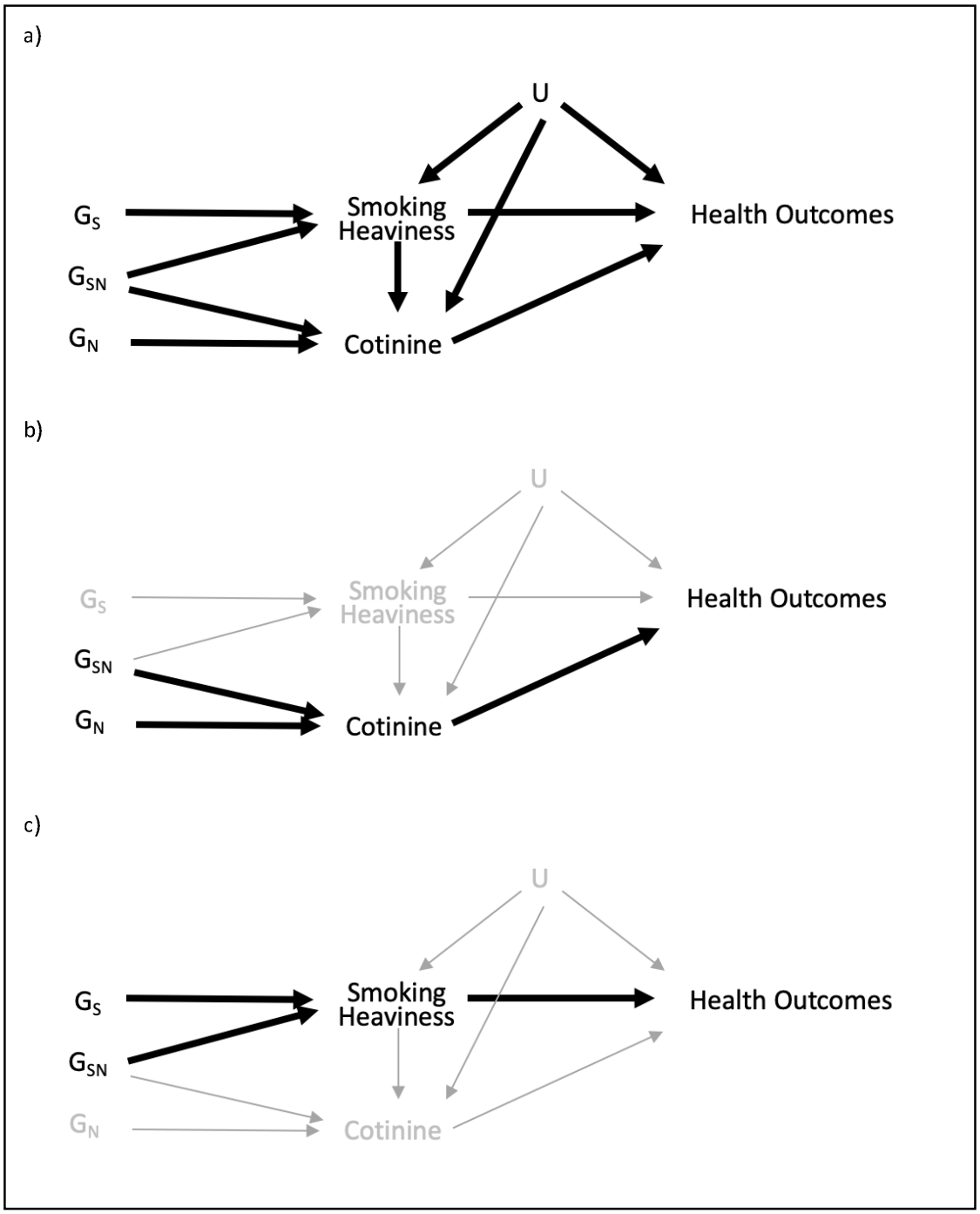
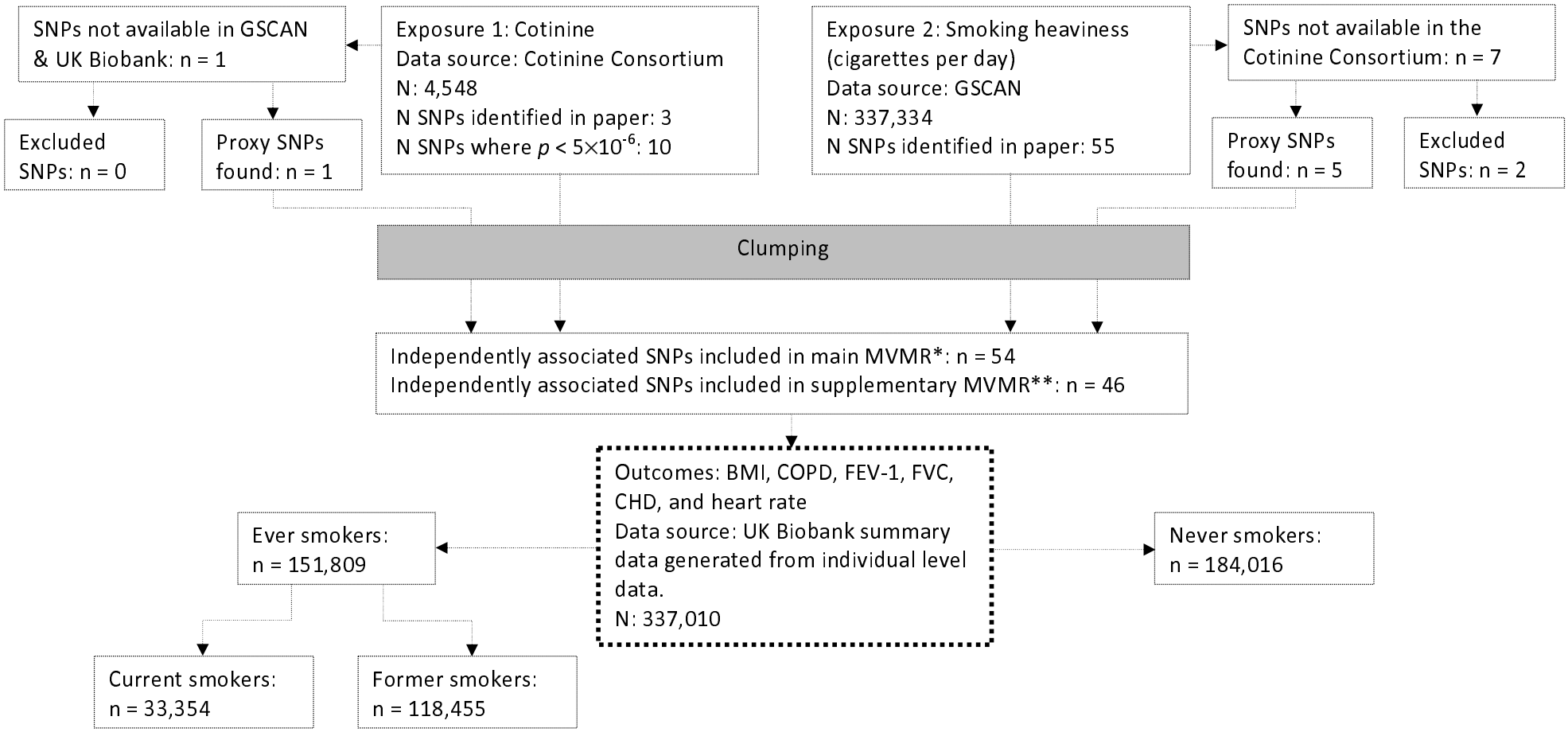


Figure 1. Directed acyclic graphs to show the relationship between genetic instruments (G_S, G_{SN}, and G_N), exposures (smoking heaviness and cotinine), confounding (U) and outcomes (health outcomes) in a multivariable Mendelian randomisation analysis.

Figure 2. A flow chart describing the inclusion process for the Multivariable Mendelian Randomisation analyses.



Note: SNPs = single nucleotide polymorphisms; BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. *SNPs were selected for inclusion in the main analysis using a *p*-value threshold ($p < 5\times10^{-6}$); ** SNPs were selected for inclusion in the supplementary analysis was completed using the genome

wide significant threshold ($p < 5\times10^{-8}$). Boxes with dashed borders indicate where individual-level data was used to generate summary-level data. Boxes with solid borders indicate summary-level data. 1,185 UK Biobank participants were excluded who preferred not to state their smoking status.

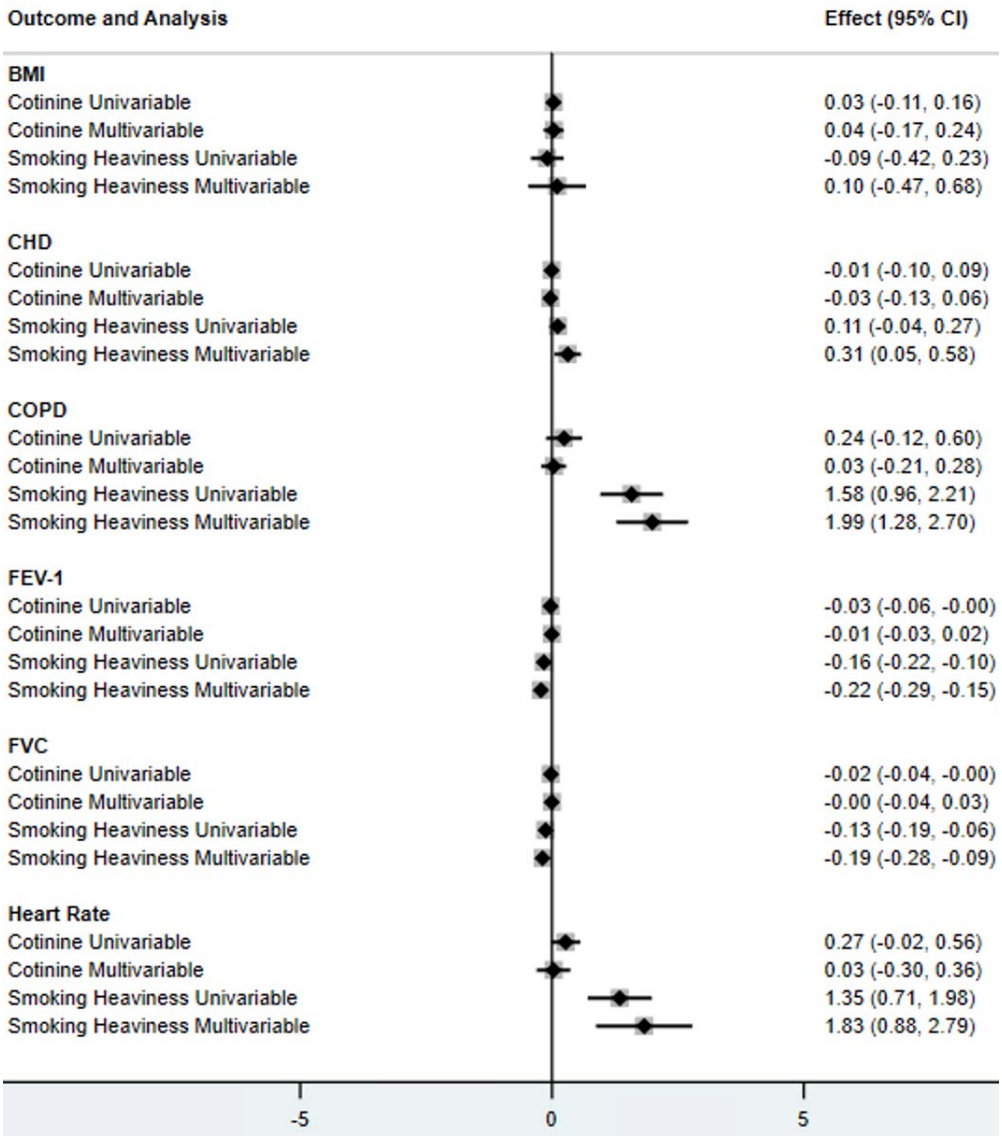


Figure 3. Univariable and multivariable Mendelian randomisation IVW analysis of cotinine and smoking heaviness (cigarettes per day) and smoking-related health outcomes among ever smokers (n = 54 SNPs).

Note: A *p*-threshold of 5×10^{-8} was used to determine the single nucleotide polymorphisms (SNPs) associated with CPD. A lower threshold of 5×10^{-6} was used to determine the SNPs associated with cotinine due to the low number of SNPs associated at the 5×10^{-8} threshold. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. Effects are betas for continuous variables (BMI, FEV-1, FVC and HR) and log odds ratios for binary outcomes (COPD and CHD) per standard deviation increase in cotinine levels/number of cigarettes per day. Univariable analyses presented are the total effects using the inverse variance weighted (IVW) method. The multivariable analyses reflect the direct effects using the MVMR-IVW method.

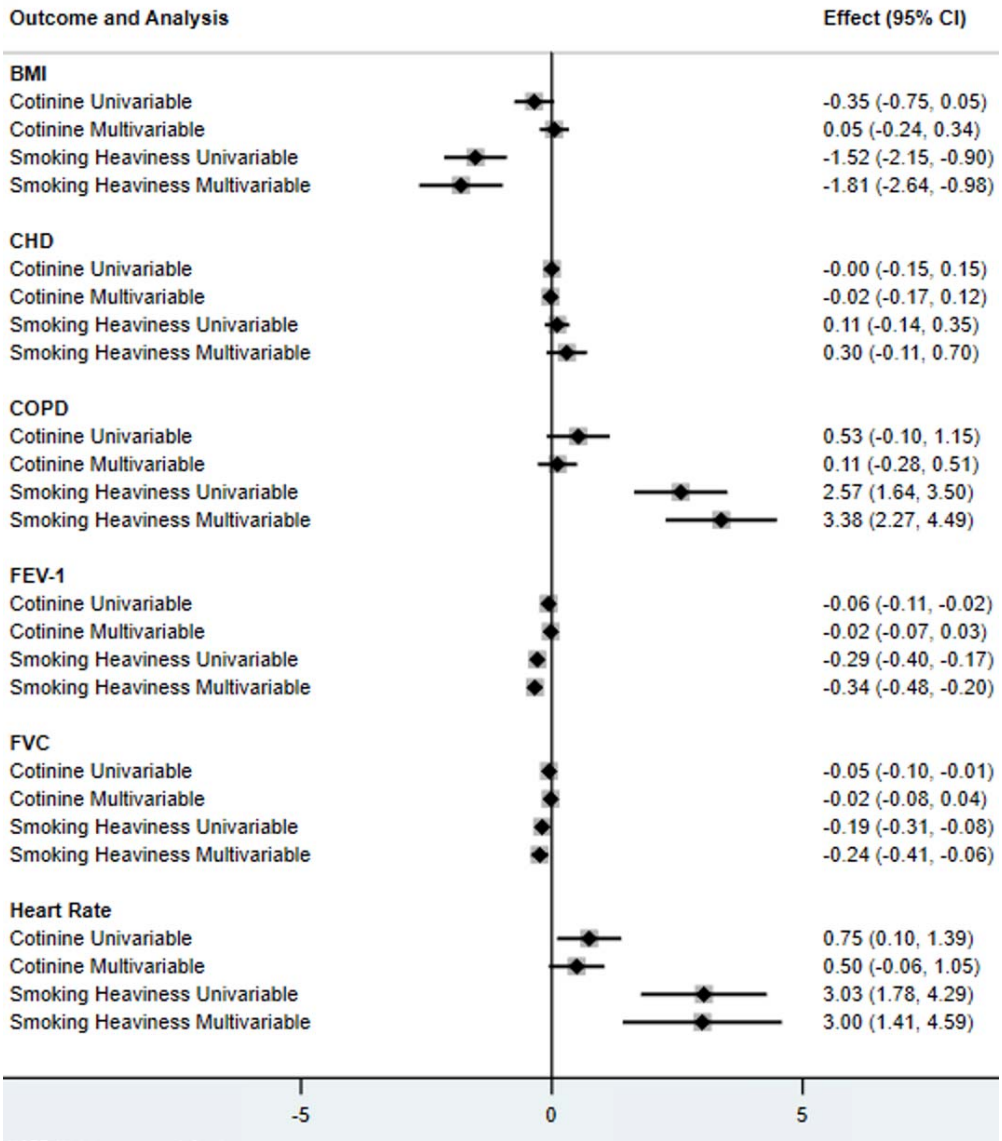


Figure 4. Univariable and multivariable Mendelian randomisation IVW analysis of cotinine and smoking heaviness (cigarettes per day) and smoking-related health outcomes among current smokers (n = 54 SNPs).

Note: A *p*-threshold of 5×10^{-8} was used to determine the single nucleotide polymorphisms (SNPs) associated with CPD. A lower threshold of 5×10^{-6} was used to determine the SNPs associated with cotinine due to the low number of SNPs associated at the 5×10^{-8} threshold. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. Effects are betas for continuous variables (BMI, FEV-1, FVC and heart rate) and log odds ratios for binary outcomes (COPD and CHD) per standard deviation increase in cotinine levels/number of cigarettes per day. Univariable analyses presented are the total effects using the inverse variance weighted (IVW) method. The multivariable analyses reflect the direct effects using the MVMR-IVW method.