Estimating the long-term health impact of nicotine exposure by dissecting the effects of nicotine versus non-nicotine constituents of tobacco smoke: A multivariable Mendelian randomisation study

Khouja, J. N.[1,2*], Sanderson, E.[2,3], Wootton, R.E.[1,2,6], Taylor, A. E.[2,3], Church, B. A.[5], Richmond, R. C.[2,3] & Munafo, M. R.[1,2,4]

[1] School of Psychological Science, University of Bristol, Bristol, United Kingdom, BS8 1TU

[2] Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK, BS8 2BN

[3] Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK, BS8 2BN

[4] NIHR Bristol Biomedical Research Centre, Bristol, UK, BS8 2BN

[5] School of Psychology and Vision Sciences, University of Leicester, UK, LE1 7HA

[6] Nic Waals Institute, Lovisenberg diakonale sykehus, Oslo, Norway, 0456

*Corresponding author

Email: jasmine.khouja@bristol.ac.uk

Short title: Health effects of nicotine versus non-nicotine constituents of tobacco smoke

Abstract

The detrimental health effects of smoking are well-known, but the impact of long-term nicotine use without exposure to the other constituents of tobacco is less clear. Given the increasing long-term use of alternative nicotine delivery systems, such as e-cigarettes, it is increasingly important to understand and separate the effects of long-term nicotine use from the impact of tobacco smoke exposure. Using a multivariable Mendelian randomisation framework, we explored the direct effects of nicotine compared with the non-nicotine constituents of tobacco smoke on health outcomes (lung cancer, chronic obstructive pulmonary disease [COPD], forced expiratory volume in one second [FEV-1], forced vital capacity [FVC], coronary heart disease [CHD], and heart rate [HR]). The results suggest that these smoking-related outcomes are not due to nicotine exposure but are caused by the other components of tobacco smoke; however, there are multiple potential sources of bias, and the results should be triangulated using evidence from a range of methodologies.

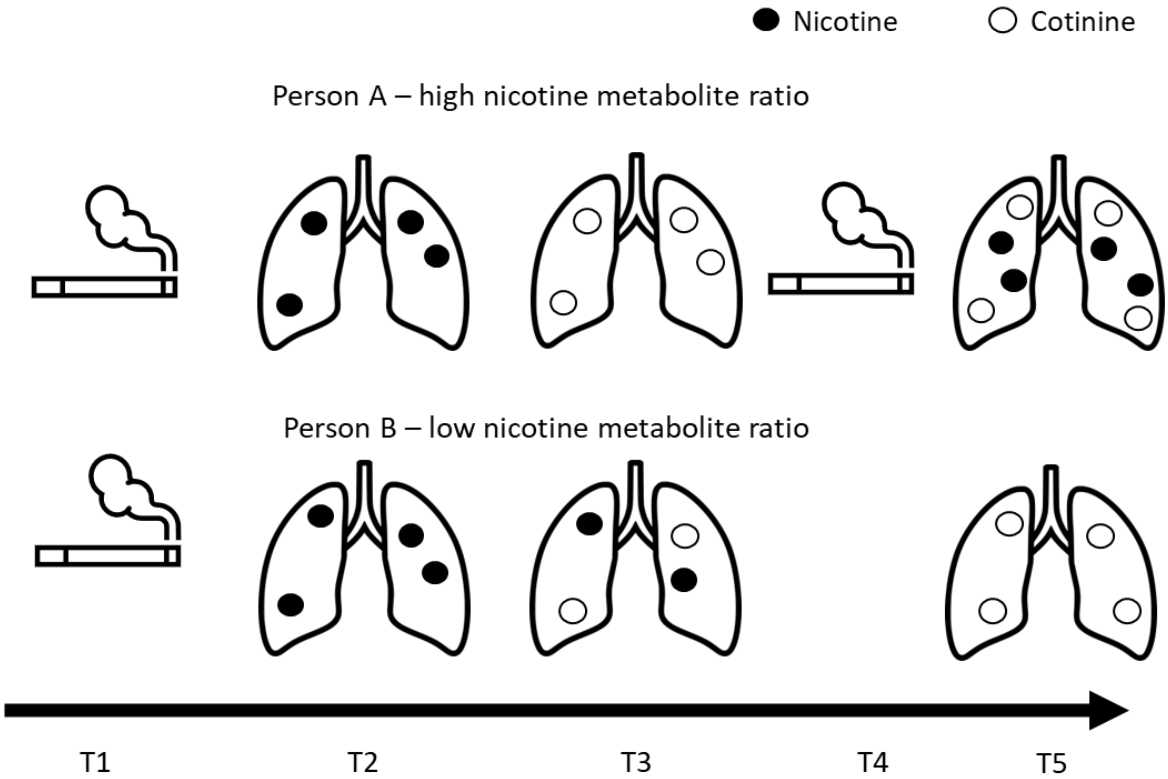
38 Although the detrimental health effects of smoking are well known [1], it is less clear which
39 individual components of cigarette smoke – such as nicotine – drive these effects. In recent
40 years, it has become increasingly important to fill this knowledge gap, given the increase in
41 the popularity of alternative nicotine delivery systems, in particular e-cigarettes [2]. E-
42 cigarettes are often used by smokers to stop smoking and are generally used for longer
43 periods than conventional cessation aids which also contain nicotine (e.g., nicotine patches)
44 [3, 4]. Although *short-term* use of nicotine without the remaining constituents of tobacco
45 smoke is relatively harmless, the *long-term* harms of nicotine use are poorly understood [5].
46 Given many smokers choose to use e-cigarettes to reduce their risk of poor health outcomes
47 [2], and that e-cigarettes can be used with or without nicotine, smokers and e-cigarette
48 users need causal evidence regarding the possible health outcomes related to long-term
49 nicotine use to make informed decisions regarding their e-cigarette use.
50
51 Randomised controlled trials (RCT) are often employed to explore causal relationships [6],
52 but an RCT would be unethical and impractical in this scenario – non-smokers would need to
53 be unnecessarily exposed to nicotine for decades to understand the (potentially harmful)
54 long-term effects of nicotine use without confounding from exposure to cigarette smoke.
55 Mendelian randomisation (MR) is an alternative method which can be used in scenarios
56 where RCTs are implausible or unethical. MR uses genetic variants associated with the
57 exposure of interest to serve as instrumental variables to estimate the effect of that
58 exposure on a particular outcome [7, 8]. Nonetheless, they require well powered genome-
59 wide association studies (GWAS) to identify such genetic variants. There are currently no
60 available GWAS of e-cigarette use or nicotine exposure without exposure to tobacco.

61 Furthermore, there are insufficient data available to conduct a well-powered GWAS. Novel
62 methods are therefore required to explore this research question.
63
64 Multivariable MR (MVMR) is an extension of MR that can be used to explore the direct
65 effect of one exposure on an outcome while accounting for the effect of another
66 (potentially correlated) exposure or exposures [9]. The method is robust to pleiotropic
67 effects of the single nucleotide polymorphisms (SNPs) used as instruments through the
68 other exposures included in the model. As large-scale GWAS have identified SNPs that are
69 associated with smoking heaviness [10] (i.e., the number of cigarettes a person smokes per
70 day) and the nicotine metabolite ratio (i.e., how quickly a person metabolises nicotine – see
71 below) [11] among smokers, we can employ this method to explore the effect of nicotine on
72 selected health outcomes. Rather than exploring the total effect of smoking on a health
73 outcome (which includes the effect of nicotine), this method allows us to separate the direct
74 effect of nicotine versus the other constituents of tobacco smoke.
75
76 However, direct measurement of nicotine is difficult. Metabolites of nicotine such as
77 cotinine (a direct metabolite) and 3'hydroxycotinine (3HC; a metabolite of cotinine) are
78 often used as objective measures to proxy for nicotine given the short half-life of nicotine
79 itself [12] (Supplementary Figure S1). Cotinine and cotinine plus 3HC are highly specific
80 biomarkers, but both are impacted by metabolism (which can differ between individuals).
81 The nicotine metabolite ratio (NMR), is a measure of how quickly a person metabolises
82 nicotine (3'hydroxycotinine/cotinine) and therefore causally impacts the amount of nicotine
83 in a person's body given a set amount of nicotine exposure [13]. Figure 1 illustrates that a
84 smoker with a higher NMR (Person A) will have less circulating nicotine in their body than a

smoker with a lower NMR (Person B) given the same level of nicotine exposure; however, because smokers with a higher NMR clear nicotine quickly from their system, this often results in them smoking more cigarettes per day (CPD) [14]. In a standard MR model with NMR as an exposure, the result may be ambiguous as it could reflect the effect of higher nicotine exposure resulting from increased smoking heaviness or lower nicotine exposure resulting from faster metabolism. In contrast, when both NMR and CPD are included in an MVMR model (Figure 2), the analysis is analogous to a fixed access experiment in which the amount of smoke exposure / cigarettes smoked is fixed and the effect of nicotine exposure per cigarette smoked can be assessed. Therefore, in an MVMR framework accounting for genetic predisposition to CPD, the results should dissect the direct effect of nicotine levels per cigarette smoked from the direct effect of the other constituents of tobacco smoke, whereby an increase in NMR reflects lower exposure to nicotine (Figure 1).

99 Figure 1. Illustration of the impact of the nicotine metabolite ratio on circulating nicotine

100 and smoking heaviness.



101

102 Note: This illustration shows the differences in nicotine exposure between two people who

103 smoke: Person A who has a high nicotine metabolite ratio, and Person B who has a low

104 nicotine metabolite ratio. At timepoint 1 (T1) both Person A and B smoke one cigarette and

105 inhale the same amount of nicotine which can be seen circulating in their body at timepoint

106 2 (T2). Later, at timepoint 3 (T3), Person A will have less circulating nicotine in their body

107 than Person B (despite having inhaled the same level of nicotine) as more of the nicotine

108 has been metabolised into cotinine. However, because smokers with higher nicotine

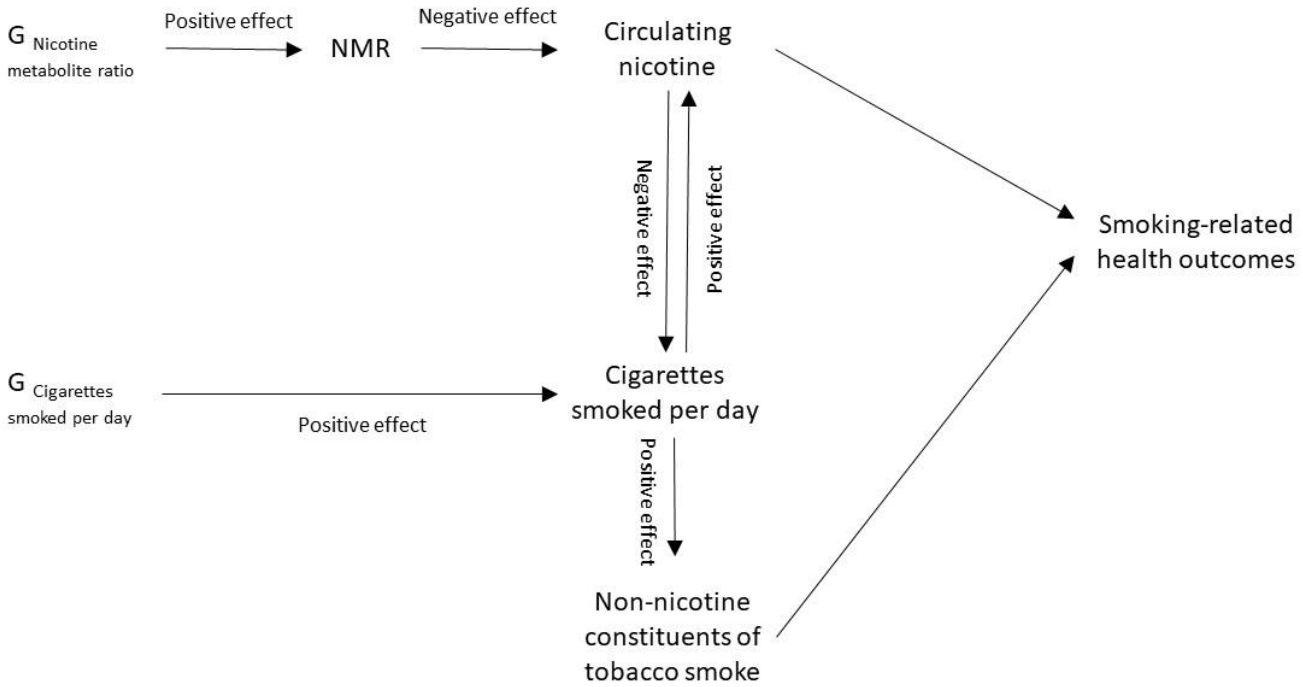
109 metabolite ratios clear nicotine more quickly from their system, this often results in them

110 smoking more cigarettes per day. So, at timepoint 4 (T4), Person A smokes another cigarette

111 whereas Person B does not. This results in Person A having more circulating nicotine than

112 Person B over the same time period [14].

113 Figure 2. Schematic of the study model



114

115 Note: This schematic shows the proposed causal pathways of the study model. Where the causal directions between variables are known

116 (i.e., evidenced in previous research), the direction of the effect is indicated as ‘positive’ or ‘negative’. NMR = Nicotine Metabolite Ratio. G =

117 genetic variants associated with the named exposure.

118 Using this framework, it is possible to explore the effects of nicotine and the effects of the
119 remaining constituents of tobacco smoke on health outcomes known to be impacted by
120 smoking, and which therefore may be impacted by long-term nicotine use, such as cancer
121 and heart and lung function [15-17]. We employed a two-sample MVMR framework to
122 explore the direct effects of nicotine compared with the non-nicotine constituents of
123 tobacco smoke on health outcomes (lung cancer, chronic obstructive pulmonary disease
124 [COPD], forced expiratory volume in one second [FEV-1], forced vital capacity [FVC],
125 coronary heart disease [CHD] and heart rate [HR]). These outcomes include positive and
126 negative control outcomes to aid causal inference; nicotine use is known to increase heart
127 rate, but is not thought to be carcinogenic, so if the results do not provide evidence of an
128 effect of nicotine on heart rate, and / or if they suggest an effect of nicotine on lung cancer,
129 the results may indicate bias from a violation of the assumptions of MR [18-20]. In
130 supplemental analyses, our secondary aim is to compare the results with analyses using
131 alternative proxies for nicotine exposure (i.e., cotinine and cotinine plus 3HC).
132

133 **Methods**

134 **Data Sources**

135 To identify relevant instrumental variables to include in the model, we sourced summary-
136 level genetic data from well-powered, published GWAS. We identified SNPs associated with
137 NMR, cotinine, cotinine plus 3HC and smoking heaviness (measured by CPD) [10, 11, 21].
138 For the outcomes, we required data to be stratified by smoking status in order to be
139 comparable with the population of the exposure GWAS data. An appropriately stratified
140 GWAS (stratified by smoking status: ever versus never [22]) was only available for one of the
141 outcomes, lung cancer. For the six remaining outcomes, we were not aware of any existing

142 available GWAS summary-level data that appropriately stratified by smoking status. We
143 therefore used UK Biobank data to conduct GWAS of CHD, COPD, FEV, FVC and HR, stratified
144 by smoking status and restricted to individuals of European ancestry. These analyses were:
145 1) stratified by whether participants had ever smoked, and further stratified into 2) current
146 smoker, or 3) former smoker, with 4) never smoker as the comparator, and were conducted
147 using the MRC Integrative Epidemiology Unit UK Biobank GWAS pipeline (version 2) [23, 24].
148 We also conducted stratified GWAS of body mass index (BMI) using UK Biobank data with
149 the intention of including BMI as an outcome. However, these methods and results are not
150 presented here as we identified a potential interpretation issue: the NMR GWAS was
151 adjusted for BMI, meaning these results would likely be biased [25].
152
153 *Smoking heaviness*
154 Liu and colleagues [10] report summary-level statistics from a GWAS meta-analysis of
155 smoking heaviness (measured by standard deviation change in CPD categories, equivalent to
156 2-3 additional cigarettes per day) among 337,334 smokers of European ancestry in the
157 GWAS and Sequencing Consortium of Alcohol and Nicotine (GSCAN). Smokers were defined
158 as having ever or currently smoked and smoking heaviness (at the time of smoking) was
159 binned into five categories or included as pre-defined bins from the original study (see
160 supplementary material in Liu and colleagues, [10]. Analyses were adjusted for age, age-
161 squared, sex, genetic principal components, and smoking status (current versus former).
162 SNPs associated with CPD explained 4% of variance in CPD. SNPs were reported as
163 independent if they explain additional variance in conditional analyses using a partial
164 correlation-based score statistic [26]. The reported data are available
165 at: https://doi.org/10.13020/3b1n-ff32.

166

167 *Nicotine*

168 Buchwald and colleagues [11] report summary-level statistics from a GWAS meta-analysis of
169 the standard deviation change in NMR among 5,185 current smokers of European descent
170 with cotinine levels ≥10 ng/ml (indicating recent smoking). NMR is a ratio of
171 3'hydroxycotinine/cotinine which indicates how quickly a person metabolises and clears
172 nicotine. Analyses were adjusted for population substructure, age, sex, BMI, alcohol use,
173 and birth year. The SNPs associated with NMR explained ~38% of the variance in NMR. SNPs
174 were reported as independent if they explained additional variance in a step-wise
175 conditional regression using genome-wide complex trait analysis [27]. The reported data are
176 available on request from the authors. Information relating to the nicotine measures used in
177 the supplemental analysis (using cotinine and cotinine plus 3HC data as a proxy for nicotine
178 exposure) can be found in the Supplementary Material (Supplemental Note S1).

179

180 *Lung Cancer*

181 McKay and colleagues [22] report summary-level statistics from a case-control GWAS meta-
182 analysis of overall lung cancer risk (including adenocarcinoma, squamous cell carcinoma,
183 and small cell carcinoma) among 50,046 individuals of European descent using the
184 International Lung Cancer Consortium (ILCCO). Analyses were adjusted for age, gender,
185 country (if applicable) and significant principal components. We received the data via
186 request to the authors and we additionally received the data stratified by smoking status
187 (ever versus never smoking). Definitions of ever and never smoking differed by cohort
188 included in the meta-GWAS. Details of the definitions can be found in McKay and
189 colleagues' [22] supplementary note.

190

191 *Non-Cancer Smoking-Related Health Outcomes*

192 We conducted GWAS of non-cancer health outcomes using data from UK Biobank, a
193 population-based health research resource consisting of approximately 500,000 people,
194 aged between 38 years and 73 years, who were recruited between the years 2006 and 2010
195 from across the UK [28]. UK Biobank has a particular focus on identifying determinants of
196 human diseases in middle-aged and older individuals who provided a wide range of health
197 information (data available at www.ukbiobank.ac.uk). A full description of the study design,
198 participants, and quality control (QC) methods have been described previously [28, 29]. UK
199 Biobank received ethics approval from the Research Ethics Committee (REC reference for UK
200 Biobank is 11/NW/0382). Written informed consent was obtained from participants prior to
201 their participation in UK Biobank. Approval to use these data was sought from and
202 approved by UK Biobank (project 9142). All studies approved by UK Biobank do not require
203 individual ethical clearance as UK Biobank has approval from the North West Multi-centre
204 Research Ethics Committee as a Research Tissue Bank under which each approved study
205 operates.

206

207 The full data release contains the cohort of successfully genotyped samples (n=488,377).
208 Further information about genotyping and imputation can be found in the Supplementary
209 Material (Supplemental Note S2). Individuals with mismatched sex were excluded and the
210 analyses were restricted to those of 'European' ancestry using an in-house k-cluster means
211 method. The GWAS were conducted using the linear mixed model (LMM) association
212 method as implemented in BOLT-LMM (v2.3) [30], adjusting for genotype array, age and sex
213 [23]. BOLT-LMM association statistics are on the linear scale, so test statistics (betas and

214 their corresponding standard errors) relating to binary phenotypes were transformed to log
215 odds ratios and their corresponding 95% confidence intervals (95% CI) on the liability scale
216 using a Taylor transformation expansion series [31]. Each outcome was stratified by self-
217 reported participant smoking status (ever [further stratified by current or former] and
218 never), resulting in four GWAS per outcome. Smoking status was categorised as never,
219 previous, and current smoking in UK Biobank (field ID 20116). From this variable, we derived
220 an 'ever smokers' category which was defined as currently or having previously smoked
221 occasionally, most days or daily (i.e., having smoked more than just once or twice). Current
222 smoking was defined as currently smoking occasionally, most days or daily. Former smoking
223 was defined as not currently smoking but having previously smoked occasionally, most days
224 or daily (i.e., more than just once or twice). Those who had tried smoking once or twice or
225 who had never smoked were categorised as never smokers.
226
227 We identified COPD cases as participants who self-reported a doctor's diagnosis of COPD.
228 Lung function (FEV-1 and FVC, in litres) was measured using a Vitalograph spirometer. CHD
229 diagnosis was determined using linked hospital admission data (ICD codes relating to
230 Ischemic Heart Disease). Further information regarding each health outcome (including UK
231 Biobank field IDs) can be found in the Supplementary Material (Supplemental Note S3).
232
233 **Statistical Analysis**
234 All analyses were conducted using R version 1.4.1.
235
236 *Selection of genetic variants*

237 We selected conditionally independent (at the genome-wide significant level, $p < 5 \times 10^{-8}$)
238 genetic variants identified in the GWAS of NMR and CPD for inclusion in the analysis [10,
239 11]; 55 SNPs associated with smoking heaviness were identified as conditionally
240 independent [10]; 7 SNPs associated with the NMR were identified as conditionally
241 independent.
242
243 After removing SNPs which were not available in the outcome GWAS (where no available
244 proxy could be identified), harmonising, and clumping the combined exposure datasets with
245 the outcome dataset, the number of SNPs in each analysis varied. Supplementary Tables S1-
246 S6 detail which SNPs were included and excluded (with reasons) in each analysis, and
247 Supplemental Note S4 gives further details of the proxy search and clumping methods.
248
249 We tested instrument strength and validity using the two-sample conditional F-statistic for
250 MVMR and Cochran *Q* statistic [9, 32]. The conditional F-statistic for MVMR indicates
251 instrument strength of each exposure when accounting for the prediction of other
252 exposures in the model (i.e., whether the SNPs jointly predict smoking heaviness after
253 predicting the NMR, and vice versa) [33]. The standard critical values for the F-statistic and
254 Q-statistic can be approximated so F-statistic should be greater than 10 to indicate sufficient
255 instrument strength and *Q* estimates should be less than the number of SNPs included in
256 the model to indicate no excessive heterogeneity [34, 35].
257
258 *Univariable Mendelian randomisation*
259 For comparison with the MVMR analysis, we considered the total effect of both the NMR
260 and smoking heaviness on each health outcome using MR. We used four complementary

261 MR methods (inverse variance weighted [MR-IVW], MR-Egger, weighted median, weighted
262 mode). Details of the univariable MR analysis methods can be found in Supplemental Note
263 S5.
264
265 *Multivariable Mendelian randomisation*
266 We explored the direct effects of both the NMR and smoking heaviness on health outcomes
267 using MVMR. We used summary data from Buchwald and colleagues' [11], GSCAN [10] and
268 ILCCO [22], and UK Biobank GWAS pipeline results. We repeated these analyses using two
269 complementary methods – MVMR-IVW and MVMR-Egger [9, 36].
270
271 Given that the GWAS of smoking heaviness was restricted to ever smokers and the GWAS of
272 the NMR was restricted to current smokers, the analyses were restricted to ever and
273 current smokers as the summary statistics are not applicable to never smokers. However,
274 the ILCCO lung cancer data have only been stratified by ever and never smoking status;
275 therefore, analyses restricted to current smokers were not possible when exploring lung
276 cancer incidence. In supplementary analyses, we additionally stratified the analysis by never
277 smokers to explore potential horizontal pleiotropy – effects observed among never smokers
278 could indicate horizontally pleiotropic effects (i.e., the included SNPs influencing the
279 outcome directly, or via another phenotype, but not through the measured exposure),
280 misreporting of smoking status, or residual population stratification. Horizontally pleiotropic
281 genetic variants are not valid instruments in MR analyses and their inclusion would result in
282 a violation of the exclusion restriction assumption (further details on the assumptions of MR
283 can be found in Supplemental Note S6). Pleiotropy robust methods (e.g., MR-Egger) are
284 robust to horizontal pleiotropy under assumptions of the form that pleiotropy takes. For all

285 outcomes except lung cancer (where data are unavailable), we also included supplementary
286 results restricted to former smokers to explore whether any health effects found among
287 current smokers may be recoverable (i.e., not present among former smokers who have
288 recovered following smoking cessation).
289
290 **Results**
291 **Descriptive Statistics**
292 Smoking heaviness data were collected from 337,334 current and former smokers who
293 reported their current or past smoking behaviour respectively (analyses corrected for
294 current versus former status, average bin 3 [SD = 1] equating to 16-25 cigarettes per day)
295 [10]. NMR data were collected from 5,185 current smokers [11]. Case rates and means (with
296 standard deviations) for the outcomes in each of the samples included in the analyses are
297 shown in Table 1.
298
299 Table 1. Descriptive statistics for participants included in the genome-wide association
300 studies of lung cancer, coronary heart disease, chronic obstructive pulmonary disease, lung
301 function and heart rate.

	Ever smokers	**Current Smokers**	**Former Smokers**	**Never Smokers**
	International Lung Cancer Consortium			
N	40,187	N/A	N/A	9,859
Lung Cancer (case rate)	57%	N/A	N/A	24%
	UK Biobank			
N	213,341	49,721	163,620	258,056
Coronary Heart Disease (case rate)	11%	11%	11%	6%
Chronic Obstructive Pulmonary Disease (case rate)	3%	6%	2%	<1%

Forced Expiratory Volume (mean litres [SD])	2.82 (0.78)	2.78 (0.83)	2.83 (0.76)	2.88 (0.78)
Forced Vital Capacity (mean litres [SD])	3.78 (0.97)	3.81 (1.02)	3.78 (0.96)	3.78 (0.99)
Heart Rate (mean beats per minute [SD])	69.05 (11.38)	71.27 (11.58)	68.40 (11.23)	68.93 (11.10)

Note: SD = standard deviation

Instrument strength and heterogeneity

Where the Cochran's Q statistic was greater than the number of SNPs included in the model, it is advised to focus on pleiotropy robust methods as this indicates heterogeneity and potential pleiotropy. MR-Egger and MVMR-Egger give estimates that are robust to directional horizontal pleiotropy under the assumption that this pleiotropy is uncorrelated with the strength of association between the SNP and the exposure [36]. However, using the MR-Egger and MVMR-Egger methods limits the statistical power of the analysis compared to MVMR-IVW. Therefore, although we present the IVW results in the text, we have explicitly stated where the sensitivity analyses differ substantially (i.e., where the results could lead to different conclusions) and in these cases, additionally compare the Egger results (and weighted mode and weighted median results) in the MR analyses.

The F-statistics indicate that the SNPs included in these analyses are strong instruments for assessing the direct effects of smoking heaviness while accounting for the effect of the NMR (Fs = 33.96 and 34.17) and for assessing the direct effects of the NMR while accounting for smoking heaviness (Fs = 30.17 and 49.08).

322 **Binary outcomes**

323 The total and direct effects of the NMR and smoking heaviness on lung cancer, CHD, COPD,
324 and are shown in Figure 3 (ever smokers). These results are also included in Supplementary
325 Tables S7 (lung cancer), S8 (CHD), and S9 (COPD) along with the F-statistics, Q-statistics and
326 Egger intercept and the results among current, never and formers smokers. We have
327 focussed on the results among ever smokers rather than current smokers here as receiving
328 a diagnosis of lung cancer, CHD or COPD may increase the likelihood of someone quitting
329 smoking, so the results are more relevant and interpretable among ever smokers. The
330 results among current smokers can be found in Supplemental Note S7 and Supplementary
331 Figure S2. Additionally, we have included the median and mode weighted MR sensitivity
332 analyses in Supplementary Table S10. Results are presented as odds ratios per standard
333 deviation (SD) increase in the exposure phenotype (i.e., per SD increase in the NMR or
334 cigarettes per day).

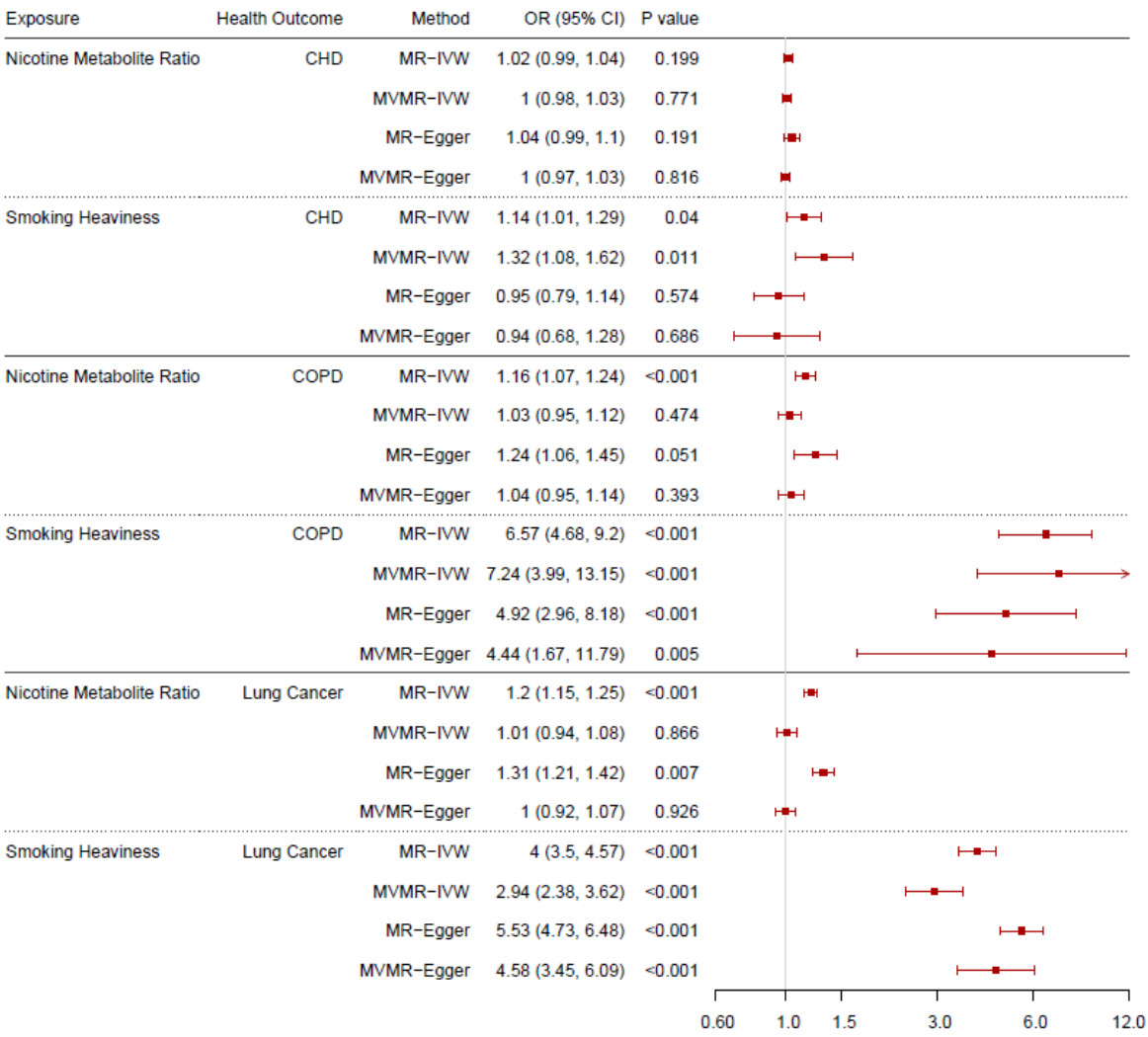


Figure 3. Forest plot displaying the effect of nicotine metabolite ratio and smoking heaviness on coronary heart disease (CHD), chronic obstructive pulmonary disease (COPD) and lung cancer incidence: univariable Mendelian randomisation (MR) and multivariable Mendelian randomisation (MVMR) results among ever smokers.

Lung Cancer

The MR-IVW results indicate strong evidence to suggest that increased NMR and increased smoking heaviness both increase the risk of developing lung cancer among ever smokers (odds ratio [OR] = 1.20, 95% confidence interval [CI] 1.15 to 1.25; OR = 4.00, 95% CI 3.50 to 4.57 respectively). However, the MVMR-IVW results do not provide clear evidence to suggest an effect of NMR on lung cancer risk (OR = 1.01, 95% CI 0.94 to 1.08), and indicate

348 that only increased smoking heaviness increased lung cancer risk (OR = 2.94, 95% CI 2.38 to
349 3.62). Although there is considerable evidence of heterogeneity and potential pleiotropy in
350 the smoking heaviness analyses among ever smokers, these results are supported by the
351 MR-Egger results. Furthermore, there is no clear evidence of pleiotropy or bias due to
352 population stratification indicated by the results among never smokers.
353
354 *Coronary Heart Disease*
355 The MR-IVW results provide no clear evidence that NMR affects CHD risk (OR = 1.02, 95% CI
356 0.99 to 1.04). This is supported by the MR-Egger results, but the weighted median and
357 weighted mode results suggested there may be a weak effect (OR = 1.03, 95% CI 1.01 to
358 1.05, OR = 1.03, 95% CI 1.01 to 1.05 respectively). The results provide some evidence that
359 increased smoking heaviness increases CHD risk (OR = 1.14, 95% CI 1.01 to 1.29) among ever
360 smokers. The weighted median and weighted mode results support this finding (OR = 1.14,
361 95% CI 1.02 to 1.29, OR = 1.14, 95% CI 1.02 to 1.27 respectively); however, the MR-Egger
362 analysis does not support this (OR = 0.95, 95% CI 0.79 to 1.14), and there is some evidence
363 of heterogeneity indicated by the Q-statistics and the Egger intercept indicated weak
364 directional pleiotropy which is supported by evidence of a protective effect among never
365 smokers. The MVMR-IVW results provide no clear evidence of an effect of NMR on CHD risk
366 (OR = 1.00, 95% CI 0.98 to 1.03) but some evidence that increased smoking heaviness
367 increases CHD risk (OR = 1.32, 95% CI 1.08 to 1.62). As with the univariable results, the
368 MVMR-Egger analyses do not support this (OR = 0.94, 95% CI 0.68 to 1.28) and there was
369 evidence of heterogeneity and potential directional pleiotropy and some evidence of
370 horizontal pleiotropy or bias due to population stratification in the analysis among never
371 smokers.

372

373 *Chronic Obstructive Pulmonary Disease*

374 The MR-IVW results indicate that increased NMR and smoking heaviness increase the risk of
375 developing COPD among ever smokers (OR = 1.16, 95% CI 1.07 to 1.24, OR = 6.57, 95% CI
376 4.68 to 9.20 respectively). However, the MVMR-IVW results indicate no clear effect of NMR
377 on COPD risk (OR = 1.03, 95% CI 0.95 to 1.12) but suggest that increased smoking heaviness
378 increases COPD risk among ever smokers (OR = 7.24, 95% CI 3.99 to 13.15). These results
379 are supported by the MR sensitivity analyses and MVMR-Egger results and there is no clear
380 evidence of heterogeneity or directional pleiotropy or horizontal pleiotropy or bias due to
381 population stratification among never smokers (where precise null effects were observed).

382

383 **Continuous outcomes**

384 The total and direct effects of the NMR and smoking heaviness on lung function (FEV-1 and
385 FVC) and heart rate are shown in Figures 4 (ever smokers) and 5 (current smokers). These
386 results are also displayed in Supplementary Tables S11 (FEV-1), S12 (FVC), and S13 (heart
387 rate) along with the F-statistics, Q-statistics and Egger intercept and the results among ever,
388 current, never and formers smokers. Additionally, we have included the median and mode
389 weighted MR sensitivity analyses in Supplementary Table 10. Results are presented as betas
390 per standard deviation (SD) increase in the exposure phenotype (i.e., per SD increase in the
391 NMR or cigarettes per day). As per MR-STROBE guidelines [37], we have reported the results
392 in text on an interpretable scale (i.e., difference in outcome in relevant units e.g., mL).
393 However, to aid comparability across outcomes, we present the results per standard
394 deviation change in the forest plots.

395

396

397

398

Exposure	Health Outcome	Method	B (95% CI)	P value
Nicotine Metabolite Ratio	FEV−1	MR−IVW	−0.02 (−0.02, −0.01)	<0.001
		MVMR−IVW	0 (−0.02, 0.01)	0.399
		MR−Egger	−0.03 (−0.04, −0.01)	0.013
		MVMR−Egger	0 (−0.01, 0.01)	0.653
Smoking Heaviness	FEV−1	MR−IVW	−0.24 (−0.28, −0.19)	<0.001
		MVMR−IVW	−0.23 (−0.31, −0.15)	<0.001
		MR−Egger	−0.25 (−0.32, −0.18)	<0.001
		MVMR−Egger	−0.23 (−0.36, −0.1)	0.001
Nicotine Metabolite Ratio	FVC	MR−IVW	−0.01 (−0.02, −0.01)	<0.001
		MVMR−IVW	−0.01 (−0.02, 0)	0.27
		MR−Egger	−0.02 (−0.03, −0.01)	0.009
		MVMR−Egger	−0.01 (−0.02, 0.01)	0.407
Smoking Heaviness	FVC	MR−IVW	−0.13 (−0.18, −0.09)	<0.001
		MVMR−IVW	−0.15 (−0.23, −0.07)	0.001
		MR−Egger	−0.1 (−0.17, −0.03)	0.008
		MVMR−Egger	−0.09 (−0.22, 0.04)	0.182
Nicotine Metabolite Ratio	HR	MR−IVW	0.01 (0, 0.01)	0.217
		MVMR−IVW	0 (−0.01, 0.01)	0.577
		MR−Egger	0 (−0.02, 0.02)	0.695
		MVMR−Egger	0 (−0.01, 0.01)	0.737
Smoking Heaviness	HR	MR−IVW	0.14 (0.1, 0.18)	<0.001
		MVMR−IVW	0.14 (0.06, 0.22)	0.001
		MR−Egger	0.11 (0.05, 0.18)	0.001
		MVMR−Egger	0.11 (−0.02, 0.24)	0.101

399

400 Figure 4. Forest plot displaying the effect of nicotine metabolite ratio and smoking
401 heaviness on standard deviations of lung function (FEV-1 and FVC), and heart rate (HR):
402 univariable Mendelian randomisation (MR) and multivariable Mendelian randomisation
403 (MVMR) results among ever smokers.

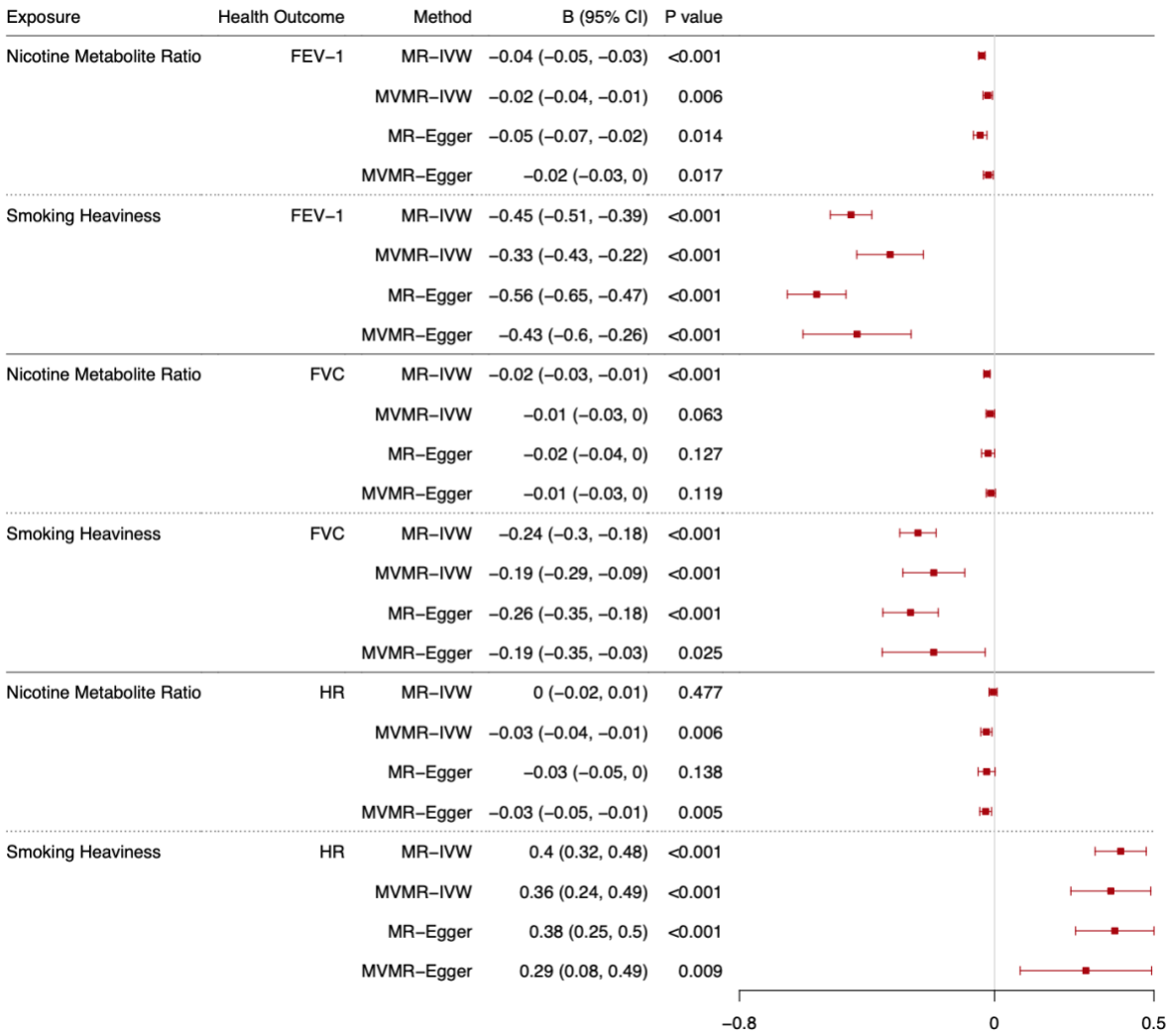


Figure 5. Forest plot displaying the effect of nicotine metabolite ratio and smoking heaviness on standard deviations of lung function (FEV-1 and FVC), and heart rate (HR): univariable Mendelian randomisation (MR) and multivariable Mendelian randomisation (MVMR) results among current smokers.

Forced Expiratory Volume in 1 Second

The results relating to FEV-1 are presented in the text as changes in millilitres (mL) of expiration per standard deviation increase in the exposure. The MR-IVW results indicate that increased NMR and smoking heaviness both decrease FEV-1 among ever smokers (Figure 4, b = -13.35, 95% CI -18.38 to -8.33, b = -185.49, 95% CI -220.39 to -150.60 respectively). However, there is no clear evidence of an effect of NMR on FEV-1 in the

417 MVMR-IVW analysis (b = 3.82, 95% CI -12.63 to 4.98), whereas there is evidence to suggest
418 increased smoking heaviness decreased FEV-1 among ever smokers (b = -176.78, 95% CI -
419 239.15 to -114.41). These results are supported by the MR sensitivity analyses and MVMR-
420 Egger results and there is no evidence of directional pleiotropy in the analyses, but there is
421 evidence of heterogeneity in all analyses except for the NMR MR analyses. However, there
422 is evidence of horizontal pleiotropy or bias due to population stratification in the NMR MR
423 analyses among never smokers.
424
425 Among current smokers (Figure 5), the MR-IVW results indicate that increased NMR and
426 smoking heaviness decreased FEV-1 (b = -33.33, 95% CI -41.76 to -24.90, b = -374.28, 95% CI
427 -428.21 to -320.35 respectively). However, the MVMR-IVW results indicate a protective
428 effect of nicotine exposure (i.e., lower nicotine exposure, due to higher NMR, lowers lung
429 function) when smoking heaviness is accounted for among current smokers (NMR b = -
430 17.77, 95% CI -29.92 to -5.62). The MVMR-IVW analyses support the MR-IVW results for
431 smoking heaviness (b = -272.36, 95% CI -359.31 to -185.42). These results are supported by
432 the MR sensitivity analyses and MVMR-Egger analyses. There is some evidence of
433 heterogeneity in all analyses except the NMR MR analyses, directional pleiotropy in the MR
434 analysis of smoking heaviness, and horizontal pleiotropy or bias due to population
435 stratification in the NMR MR analyses among never smokers.
436
437 *Forced Vital Capacity*
438 The results relating to FVC are presented in the text as changes in mL of capacity per
439 standard deviation increase in the exposure. The MR-IVW results indicate that increased
440 NMR and smoking heaviness decreased FVC among ever smokers (Figure 4, b = -14.34, 95%

441 CI -19.42 to -9.26, b = -129.63, 95% CI -174.27 to -85.00 respectively). However, there is no
442 clear evidence of an effect of NMR on FVC in the MVMR-IVW analysis (b = -6.32, 95% CI -
443 17.41 to 4.77), whereas there is evidence to suggest increased smoking heaviness decreased
444 FVC among ever smokers, after accounting for NMR (b = -149.55, 95% CI -228.07 to -71.03).
445 These results are supported by the MR sensitivity analyses and MVMR-Egger results,
446 although there is less clear evidence of an effect in the MVMR-Egger results for smoking
447 heaviness than in the MR-Egger results. There is some evidence of heterogeneity in all
448 analyses except the NMR MR analyses, but no clear evidence of directional pleiotropy.
449 There is weak evidence to suggest potential horizontal pleiotropy / bias due to population
450 stratification in the NMR MR analyses and the smoking heaviness MR-IVW analysis among
451 never smokers.
452
453 Among current smokers (Figure 5), the MR-IVW results suggest that increased NMR and
454 smoking heaviness decreased FVC (b = -24.50, 95% CI -34.61 to -14.39, b = -246.13, 95% CI -
455 304.36 to -187.89 respectively). MVMR-IVW results support the evidence for an effect of
456 smoking cessation on FVC but indicate weak evidence of a protective effect of nicotine
457 exposure when smoking heaviness is accounted for (NMR b = -13.50, 95% CI -27.39 to 0.39).
458 The MR median and mode weighted analyses support these findings, and the MR-Egger and
459 MVMR-Egger results are in the same direction but provide weaker support for the NMR
460 results. There is evidence in the MVMR-IVW analysis to suggest that increased smoking
461 heaviness decreases FVC (b -195.31, 95% CI -294.72 to -95.90), which is supported by the
462 MVMR-Egger results. There is some evidence of heterogeneity in all analyses except the
463 NMR MR analyses, but no clear evidence of directional pleiotropy. There is weak evidence

464 to suggest potential horizontal pleiotropy or bias due to population stratification in the NMR
465 MR analyses and the smoking heaviness MR-IVW analysis among never smokers.
466
467 *Heart Rate*
468 The results relating to heart rate are presented in the text as changes in beats per minute
469 per standard deviation increase in the exposure. Among ever smokers (Figure 4), the MR-
470 IVW results suggest that increased NMR does not affect heart rate (b = 0.06, 95% CI -0.04 to
471 0.16), but increased smoking heaviness does increase heart rate (b = 1.61, 95% CI 1.12 to
472 2.10). The MVMR-IVW results support the results for both NMR and smoking heaviness (b =
473 -0.04, 95% CI -0.16 to 0.09, b = 1.63, 95% CI 0.73 to 2.54 respectively). The MR sensitivity
474 analyses and MVMR-Egger results also support these findings. There is some evidence of
475 heterogeneity in all analyses, but no clear evidence of directional pleiotropy. There is weak
476 evidence to suggest potential horizontal pleiotropy or bias due to population stratification in
477 the MR-IVW NMR, MR-IVW and MR-Egger smoking heaviness analyses among never
478 smokers.
479
480 Among current smokers (Figure 5), the MR-IVW results indicate no clear effect of NMR on
481 heart rate (b = -0.05, 95% CI -0.20 to 0.09) but indicate that increased smoking heaviness
482 increases heart rate (b = 4.58, 95% CI 3.65 to 5.51). However, the MVMR-IVW results
483 suggest that decreased NMR when accounting for smoking heaviness (i.e., increased
484 nicotine exposure per cigarette) increases heart rate among current smokers (b = -0.30, 95%
485 CI -0.50 to -0.10) and increased smoking heaviness accounting for nicotine exposure
486 increases heart rate (b = 4.22, 95% CI 2.77 to 5.68). The MR sensitivity and MVMR-Egger
487 results support the MR-IVW and MVMR-IVW findings. There is some evidence of

heterogeneity in all analyses except the MR NMR analyses, but no clear evidence of directional pleiotropy. There is weak evidence to suggest potential horizontal pleiotropy / bias due to population stratification in the MR-IVW NMR, MR-IVW and MR-Egger smoking heaviness analyses among never smokers.

Interpreting the NMR MVMR results

In the MVMR results above, we interpret an increase in the direct effect of NMR as a reduction in nicotine exposure per cigarette smoked. Supplementary Table S14 illustrates how the estimated direct effects of NMR found can be flipped to indicate the effect of increased nicotine exposure per cigarette smoked on the outcome.

Sensitivity and supplementary analysis

The results of the sensitivity analyses whereby cotinine and cotinine plus 3HC were used as alternative proxies for nicotine exposure can be found in Supplementary Tables S7-S13. The results are generally in line with the main results with some exceptions. Increased cotinine and cotinine plus 3HC exposure appears to increase risk of lung cancer (to a lesser extent than non-nicotine constituents of tobacco smoke). There is also less evidence to suggest a protective effect of nicotine on COPD among current smokers in the analyses alternatively including cotinine and cotinine plus 3HC and there is less evidence of an effect on FEV-1 and heart rate among current smokers in the analyses including cotinine plus 3HC.

The results among former smokers can be found in Supplementary Tables S7-S13. The findings suggest that there are likely lasting detrimental effects of smoking, but they are unlikely to be attributable to nicotine exposure.

512 **Discussion**

513 Overall, we found little evidence to suggest a major detrimental effect of nicotine exposure
514 on health when using an MVMR model to distinguish the direct effects of nicotine from the
515 direct effects of non-nicotine components of tobacco smoke (using NMR and CPD as
516 proxies). Our model suggests that, among current smokers, increased nicotine exposure per
517 cigarette smoked increases heart rate but may have protective effects on lung function and
518 COPD. Our findings in relation to smoking heaviness were generally as expected except for
519 CHD, for which there was limited evidence of an effect. In line with previous research [15-
520 17], increased smoking heaviness detrimentally impacted risk of lung cancer, COPD, and
521 CHD, as well as lung function and heart rate. As expected, and in line with evidence from a
522 range of human and animal studies [19], our analysis suggests that exposure to nicotine
523 without the remaining constituents of tobacco smoke increases heart rate (the positive
524 control).

525

526 In contrast to previous evidence, in which increased smoking heaviness was found to
527 increase risk of CHD, we found limited evidence of an effect of smoking heaviness on CHD
528 among ever smokers. Previous MR studies have also shown minimal evidence of this effect
529 using UK Biobank data [38], which could indicate that the outcome GWAS sample may not
530 be representative of the general population. Indeed, UK Biobank participants are more likely
531 to describe themselves as female, have fewer health issues, live in less socioeconomically
532 deprived areas, and be more highly educated, than the general population and thus the
533 results may suffer from “healthy volunteer” selection bias [39]. Given that CHD may
534 disproportionately affect men [40] and people with generally poorer health, this may

535 explain the limited findings. Additionally, we may have had limited power to detect this
536 effect given the range of causes of CHD and relatively few cases in UK Biobank.
537
538 Another unexpected finding was the apparent protective effects of nicotine on lung health
539 (COPD, FEV-1 and FVC) among current smokers. However, this finding may be due to those
540 with a lower NMR (and greater circulating nicotine per cigarette) needing to inhale less
541 smoke and other non-nicotine constituents per cigarette. Measuring smoking heaviness by
542 the number of cigarettes smoked per day will not completely capture the variation in
543 smoking intensity, so there could be a small effect of intensity which is captured in the
544 instrumental variable for NMR because increased NMR may lead to increased inhalation per
545 cigarette. Given the magnitude of effects found for NMR are much smaller than for smoking
546 heaviness, the residual variance in smoking intensity impacting the NMR estimates appears
547 to be limited. Furthermore, if there is a true protective effect it may not be clinically
548 important given the magnitude of the effects found. For example, patients only perceive
549 differences in FEV-1 with changes of 112 mL [41] whereas the MVMR-IVW analyses
550 indicated an increase of only 17.77 mL per SD increase in NMR.
551
552 In the supplementary analyses, we also saw apparent effects of cotinine and cotinine plus
553 3HC on lung cancer risk in the MVMR models. Given that nicotine is not thought to be
554 carcinogenic, these findings could indicate pleiotropic pathways are involved for these
555 additional exposures (e.g., via metabolism). This highlights that NMR may be a more
556 appropriate proxy for nicotine exposure in this context; NMR is a measure of nicotine
557 metabolism, therefore by including it in an MVMR model which is robust to pleiotropy via
558 each exposure in the model, we account for pleiotropic pathways via metabolism.

559

560 Although this is the first study to use MVMR to explore the impact of long-term nicotine
561 inhalation in humans and it may provide an early indication of the health impact of inhaling
562 nicotine via e-cigarettes, the study is not without limitations. In addition to potential issues
563 with “healthy volunteer” selection bias, the study is limited by the potential presence of
564 horizontal pleiotropy in some analyses (as indicated by effects seen in the analyses
565 restricted to non-smokers and the Egger analyses), which violates an assumption of the
566 method. The effect estimates among never smokers cannot be meaningfully interpreted as
567 we know that never smokers (or at least those who accurately self-report never smoking) do
568 not smoke any cigarettes per day despite being predisposed to heavier smoking, but
569 evidence of an effect in never smokers (along with a high Cochran’s Q statistic) is indicative
570 of horizontal pleiotropy. Despite this, no effects were observed in the relationship between
571 NMR and the health outcomes in the MVMR analyses restricted to never smokers.
572 Additionally, while the MR-Egger test of directional pleiotropy indicated directional
573 pleiotropic effects in the relationship between smoking heaviness and health outcomes,
574 there was no clear evidence of directional horizontal pleiotropic effects in the relationship
575 between NMR and health outcomes among ever and current smokers. Therefore, these
576 sensitivity analyses suggest our interpretation of the direct effects of nicotine on the
577 selected health outcomes should not be impacted by directional horizontal pleiotropy or
578 population stratification.

579

580 Further limitations of the study relate to the adjustment and inclusion criteria of the GWAS
581 used for analysis. The first issue pertains to differences between the exposure GWAS
582 inclusion criteria: the GWAS of NMR was restricted to current smokers whereas the GWAS

583 of smoking heaviness was restricted to ever regular smokers. However, smoking heaviness
584 was assessed as current or past smoking heaviness (i.e., the measure is representative of
585 the behaviour of current smokers despite being a retrospective measure, therefore we are
586 assuming that the two populations are similar). If this assumption in incorrect, we may have
587 introduced bias into our findings. Second, in contrast to the GWAS of smoking behaviour
588 and the health outcomes, the NMR GWAS was adjusted for BMI. Adjusting for BMI in one
589 exposure variable meant that we could not clearly assess the effect of nicotine on BMI,
590 however it also may impact our interpretation of results where the outcomes are associated
591 with BMI (i.e., where BMI is a plausible co-variate of the exposure and outcome
592 phenotypes) [25]. As lung and heart health could be impacted by BMI [42, 43], some of the
593 effect sizes observed may be biased towards the null. Although there are approaches
594 available to correct for this bias [44], it is not feasible to implement this correction with the
595 limited number of SNPs included in this analysis. Therefore, we must interpret these
596 findings with caution. Third, we did not use data from an updated CPD GWAS with a larger
597 sample size and more diverse sample that was released in 2022 [45] after analysis was
598 complete for this study. The study team agreed that it would be detrimental to use the new
599 data release restricted to European ancestry. The number of SNPs found to be associated
600 with smoking heaviness was three times greater in the 2022 data release compared to the
601 2019 data release, which would cause a problematic imbalance/mismatch between the
602 number of SNPs used to instrument smoking heaviness versus NMR, and potentially lead to
603 weak instrument bias [33]. Furthermore, we were unable to use data from the multi-
604 ancestry GWAS of smoking heaviness [45] as the NMR GWAS did not conduct similar
605 analyses and was restricted to those of European descent and an assumption of MVMR is
606 that all data included in an MVMR model are from the same underlying population [9].

607 Future work could focus on the inclusion of multiple ancestries if appropriate NMR GWAS
608 become available to explore whether the findings generalise to non-European ancestries.
609
610 This work highlights the limited impact that long-term nicotine exposure via inhalation (e.g.,
611 via smoking or potentially vaping) is likely to have on heart and lung health. Given that it will
612 not be possible to explore the long-term effects of vaping for some years yet, and the
613 impact of long-term nicotine use is usually confounded by smoking, this is an important
614 addition to our knowledge with respect to the harms of vaping. However, there are many
615 more potential health outcomes (e.g., the effects on cancer progression, mental health,
616 sleep, perinatal outcomes etc.) which should be explored in future research. Additionally,
617 the effect of vaping nicotine could be impacted by nicotine absorption potentially occurring
618 by a different route to smoking (e.g., via oral mucosa absorption rather than lungs), or by
619 the difference in vehicle-delivery (i.e., via liquid droplets of propylene glycol and glycerol
620 rather than particulates) [46]. Furthermore, these findings cannot inform our understanding
621 of non-nicotine e-cigarette vapour constituent exposure; further research is therefore
622 needed to understand the health impact of long-term exposure to propylene glycol,
623 vegetable glycerine and common e-liquid flavourings. Nevertheless, the findings support the
624 safe long-term use of traditional pharmaceutical nicotine replacement therapies such as
625 nicotine patches.
626
627 **Conclusions**
628 In conclusion, the present findings indicate that long-term nicotine use (via inhalation) when
629 accounting for exposure to non-nicotine constituents of tobacco smoke may increase heart
630 rate but does not increase risk of COPD or CHD and does not appear to adversely impact

631 lung function. We found that, aside from effects on heart rate (which were expected given
632 our knowledge of effects of short-term nicotine use), there was no evidence to suggest that
633 long-term nicotine exposure is responsible for the detrimental effects of smoking on the
634 outcomes that were included in this analysis. Although further research is necessary to
635 explore other health outcomes and triangulate these findings, our results support existing
636 evidence which suggests nicotine use is not a major risk factor in the development of
637 smoking-related disease.
638

Acknowledgements

We would like to thank Rachel Tyndale and her team for their support in interpreting the MVMR results relating to NMR and provision of summary statistics. We would also like to thank the GSCAN and ILCCO teams (including James McKay) for the provision of summary statistics. Quality Control filtering of the UK Biobank data was conducted by R. Mitchell, G. Hemani, T. Dudding, L. Corbin, S. Harrison, L. Paternoster as described in the published protocol (doi: 10.5523/bris.1ovaau5sxunp2cv8rcy88688v). The MRC IEU UK Biobank GWAS pipeline was developed by B. Elsworth, R. Mitchell, C. Raistrick, L. Paternoster, G. Hemani, T. Gaunt (doi: 10.5523/bris.pnoat8cxo0u52p6ynfaekeigi).

Funding

JNK, RCR, MRM, ES, and AET are all members of the Medical Research Council (MRC) Integrative Epidemiology Unit at the University of Bristol (MC_UU_00032/01, MC_UU_00032/04, MC_UU_00032/07). REW is funded by the Norwegian South-Eastern Regional Health Authority (2020024). This work was supported by Cancer Research UK (grant number C18281/A29019). This work is supported by the National Institute for Health Research (NIHR) 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 NIHR or the Department of Health and Social Care.

References

1. Das SK. Harmful health effects of cigarette smoking. Mol Cell Biochem. 2003;253(1-2):159-65. Epub 2003/11/19. doi: 10.1023/a:1026024829294. PubMed PMID: 14619966.
2. Action on Smoking and Health. Use of e-cigarettes (vapes) among adults in Great Britain. 2022. Available from: https://ash.org.uk/uploads/Use-of-e-cigarettes-vapes-among-adults-in-Great-Britain-2022.pdf?v=1661865959.
3. Jackson SE, Hill E, Shahab L, Beard E, Michie S, Brown J. Prevalence and correlates of long-term e-cigarette and nicotine replacement therapy use: a prospective study in England. BMJ Open. 2019;9(10):e029252. Epub 2019/10/13. doi: 10.1136/bmjopen-2019-029252. PubMed PMID: 31604784; PubMed Central PMCID: PMCPMC6797372.
4. Shahab L, Dobbie F, Hiscock R, McNeill A, Bauld L. 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):81-8. Epub 2016/09/25. doi: 10.1093/ntr/ntw258. PubMed PMID: 27664995; PubMed Central PMCID: PMCPMC5896546.
5. Lee PN, Fariss MW. A systematic review of possible serious adverse health effects of nicotine replacement therapy. Arch Toxicol. 2017;91(4):1565-94. Epub 2016/10/05. doi: 10.1007/s00204-016-1856-y. PubMed PMID: 27699443; PubMed Central PMCID: PMCPMC5364244.
6. Gvozdenović E, Malvisi L, Cinconze E, Vansteelandt S, Nakanwagi P, Aris E, et al. Causal inference concepts applied to three observational studies in the context of vaccine development: from theory to practice. BMC Medical Research Methodology. 2021;21(1):35. doi: 10.1186/s12874-021-01220-1.
7. Richmond RC, Davey Smith G. Mendelian Randomization: Concepts and Scope. Cold Spring Harb Perspect Med. 2022;12(1). Epub 2021/08/25. doi: 10.1101/cshperspect.a040501. PubMed PMID: 34426474; PubMed Central PMCID: PMCPMC8725623.
8. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. Nature Reviews Methods Primers. 2022;2(1):6. doi: 10.1038/s43586-021-00092-5.
9. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol. 2019;48(3):713-27. Epub 2018/12/12. doi: 10.1093/ije/dyy262. PubMed PMID: 30535378; PubMed Central PMCID: PMCPMC6734942.
10. Liu MZ, 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. Nat Genet. 2019;51(2):237-+. doi: 10.1038/s41588-018-0307-5. PubMed PMID: WOS:000457314300010.
11. Buchwald J, Chenoweth MJ, Palviainen T, Zhu G, Benner C, Gordon S, et al. Genome-wide association meta-analysis of nicotine metabolism and cigarette consumption measures in smokers of European descent. Mol Psychiatry. 2020. Epub 2020/03/12. doi: 10.1038/s41380-020-0702-z. PubMed PMID: 32157176; PubMed Central PMCID: PMCPMC7483250.
12. Benowitz NL, Hukkanen J, Jacob P, 3rd. Nicotine chemistry, metabolism, kinetics and biomarkers. Handb Exp Pharmacol. 2009;(192):29-60. Epub 2009/02/03. doi: 10.1007/978-3-540-69248-5_2. PubMed PMID: 19184645; PubMed Central PMCID: PMCPMC2953858.

13. Dempsey D, Tutka P, Jacob III P, Allen F, Schoedel K, Tyndale RF, et al. Nicotine metabolite ratio as an index of cytochrome P450 2A6 metabolic activity. Clinical Pharmacology & Therapeutics. 2004;76(1):64-72. doi: https://doi.org/10.1016/j.clpt.2004.02.011.

14. Carroll DM, Murphy SE, Benowitz NL, Strasser AA, Kotlyar M, Hecht SS, et al. Relationships between the Nicotine Metabolite Ratio and a Panel of Exposure and Effect Biomarkers: Findings from Two Studies of U.S. Commercial Cigarette Smokers. Cancer Epidemiol Biomarkers Prev. 2020;29(4):871-9. Epub 2020/02/14. doi: 10.1158/1055-9965.Epi-19-0644. PubMed PMID: 32051195; PubMed Central PMCID: PMCPMC7437528.

15. West R. Tobacco smoking: Health impact, prevalence, correlates and interventions. Psychol Health. 2017;32(8):1018-36. Epub 2017/05/30. doi: 10.1080/08870446.2017.1325890. PubMed PMID: 28553727; PubMed Central PMCID: PMCPMC5490618.

16. Jayes L, Haslam PL, Gratziou CG, Powell P, Britton J, Vardavas C, et al. SmokeHaz: Systematic Reviews and Meta-analyses of the Effects of Smoking on Respiratory Health. Chest. 2016;150(1):164-79. Epub 2016/04/23. doi: 10.1016/j.chest.2016.03.060. PubMed PMID: 27102185.

17. Khan A, Shabbir K, Ansari JK, Zia N. Comparison of forced expiratory volume in one second (FEV1) among asymptomatic smokers and non-smokers. J Pak Med Assoc. 2010;60(3):209-13. Epub 2010/03/17. PubMed PMID: 20225780.

18. Hammerton G, Munafò MR. Causal inference with observational data: the need for triangulation of evidence. Psychological Medicine. 2021;51(4):563-78. Epub 2021/03/08. doi: 10.1017/S0033291720005127.

19. Benowitz NL, Burbank AD. Cardiovascular toxicity of nicotine: Implications for electronic cigarette use. Trends Cardiovasc Med. 2016;26(6):515-23. Epub 2016/04/16. doi: 10.1016/j.tcm.2016.03.001. PubMed PMID: 27079891; PubMed Central PMCID: PMCPMC4958544.

20. Royal College of Physicians. Nicotine without smoke: Tobacco harm reduction. . 2016.

21. Ware JJ, Chen X, Vink J, Loukola A, Minica C, Pool R, et al. Genome-Wide Meta-Analysis of Cotinine Levels in Cigarette Smokers Identifies Locus at 4q13.2. Sci Rep. 2016;6:20092. Epub 2016/02/03. doi: 10.1038/srep20092. PubMed PMID: 26833182; PubMed Central PMCID: PMCPMC4735517.

22. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nat Genet. 2017;49(7):1126-32. Epub 2017/06/12. doi: 10.1038/ng.3892. PubMed PMID: 28604730.

23. Elsworth B, Mitchell, R., Raistrick, C., Paternoster, L., Hemani, G., Gaunt T. MRC IEU UK Biobank GWAS pipeline, version 2, 18/01/2019. databris. 2019. doi: 10.5523/bris.pnoat8cxo0u52p6ynfaekeigi.

24. Mitchell R, Hemani G, Dudding T, Corbin L, Harrison S, Paternoster L. UK Biobank Genetic Data: MRC-IEU Quality Control, version 2 - Datasets - data.bris. databris. 2018. doi: doi:10.5523/bris.1ovaau5sxunp2cv8rcy88688v.

25. Hartwig FP, Tilling K, Davey Smith G, Lawlor DA, Borges MC. Bias in two-sample Mendelian randomization when using heritable covariable-adjusted summary associations. International Journal of Epidemiology. 2021;50(5):1639-50. doi: 10.1093/ije/dyaa266.

26. Jiang Y, Chen S, McGuire D, Chen F, Liu M, Iacono WG, 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.
27. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88(1):76-82. Epub 2010/12/17. doi: 10.1016/j.ajhg.2010.11.011. PubMed PMID: 21167468.
28. Allen NE, Sudlow C, Peakman T, Collins R, Biobank UK. UK biobank data: come and get it. Sci Transl Med. 2014;6(224):224ed4. Epub 2014/02/21. doi: 10.1126/scitranslmed.3008601. PubMed PMID: 24553384.
29. Mitchell R, Hemani G, Dudding T, Corbin L, Harrison S, Paternoster L. UK Biobank Genetic Data: MRC-IEU Quality Control, version 2. 2019. doi: https://doi.org/10.5523/bris.1ovaau5sxunp2cv8rcy88688v.
30. Loh PR, Tucker G, Bulik-Sullivan BK, Vilhjalmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet. 2015;47(3):284-90. Epub 2015/02/03. doi: 10.1038/ng.3190. PubMed PMID: 25642633; PubMed Central PMCID: PMCPMC4342297.
31. Loh PR, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale datasets. Nat Genet. 2018;50(7):906-8. doi: 10.1038/s41588-018-0144-6. PubMed PMID: WOS:000437224400002.
32. Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. Stat Med. 2021;40(25):5434-52. Epub 2021/08/03. doi: 10.1002/sim.9133. PubMed PMID: 34338327; PubMed Central PMCID: PMCPMC9479726.
33. 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.
34. Bowden J, Del Greco M F, Minelli C, Zhao Q, Lawlor DA, Sheehan NA, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. International Journal of Epidemiology. 2018;48(3):728-42. doi: 10.1093/ije/dyy258.
35. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Statistics in Medicine. 2008;27(8):1133-63. doi: https://doi.org/10.1002/sim.3034.
36. 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.
37. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. BMJ. 2021;375:n2233. doi: 10.1136/bmj.n2233.
38. King C, Mulugeta A, Nabi F, Walton R, Zhou A, Hyppönen E. Mendelian randomization case-control PheWAS in UK Biobank shows evidence of causality for smoking intensity in 28 distinct clinical conditions. EClinicalMedicine. 2020;26:100488. Epub 2020/10/23. doi: 10.1016/j.eclinm.2020.100488. PubMed PMID: 33089118; PubMed Central PMCID: PMCPMC7564324.
39. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With

797 Those of the General Population. American Journal of Epidemiology. 2017;186(9):1026-34.
798 doi: 10.1093/aje/kwx246.
799 40. Bots SH, Peters SAE, Woodward M. Sex differences in coronary heart disease and
800 stroke mortality: a global assessment of the effect of ageing between 1980 and 2010. BMJ
801 Global Health. 2017;2(2):e000298. doi: 10.1136/bmjgh-2017-000298.
802 41. Donohue JF. Minimal Clinically Important Differences in COPD Lung Function. COPD:
803 Journal of Chronic Obstructive Pulmonary Disease. 2005;2(1):111-24. doi: 10.1081/COPD-
804 200053377.
805 42. Hypponen E, Mulugeta A, Zhou A, Santhanakrishnan VK. A data-driven approach for
806 studying the role of body mass in multiple diseases: a phenome-wide registry-based case-
807 control study in the UK Biobank. Lancet Digit Health. 2019;1(3):e116-e26. Epub 2019/07/01.
808 doi: 10.1016/S2589-7500(19)30028-7. PubMed PMID: 33323262.
809 43. Peralta GP, Marcon A, Carsin A-E, Abramson MJ, Accordini S, Amaral AF, et al. Body
810 mass index and weight change are associated with adult lung function trajectories: the
811 prospective ECRHS study. Thorax. 2020;75(4):313-20. doi: 10.1136/thoraxjnl-2019-213880.
812 44. Gilbody J, Borges MC, Smith GD, Sanderson E. Multivariable MR can mitigate bias in
813 two-sample MR using covariable-adjusted summary associations. medRxiv.
814 2022:2022.07.19.22277803. doi: 10.1101/2022.07.19.22277803.
815 45. Saunders GRB, Wang X, Chen F, Jang S-K, Liu M, Wang C, et al. Genetic diversity fuels
816 gene discovery for tobacco and alcohol use. Nature. 2022;612(7941):720-4. doi:
817 10.1038/s41586-022-05477-4.
818 46. Farsalinos KE, Spyrou A, Tsimopoulou K, Stefopoulos C, Romagna G, Voudris V.
819 Nicotine absorption from electronic cigarette use: comparison between first and new-
820 generation devices. Sci Rep. 2014;4:4133. Epub 2014/02/27. doi: 10.1038/srep04133.
821 PubMed PMID: 24569565; PubMed Central PMCID: PMCPMC3935206.
822