

1 **Health behaviours prior to pregnancy and fertility outcomes: Triangulation of evidence in**  
2 **the Norwegian Mother, Father and Child Cohort Study (MoBa)**

3  
4 **Authors**

5 Robyn E Wootton<sup>1,2,3</sup>, Rebecca B Lawn<sup>4</sup>, Maria Magnus<sup>5</sup>, Jorien Treur<sup>6</sup>, Elizabeth Corfield<sup>1,7</sup>, Pål R.  
6 Njølstad<sup>8,9</sup>, Ole A. Andreassen<sup>10</sup>, Deborah A. Lawlor<sup>2,3</sup>, Marcus R Munafò<sup>2,11</sup>, Siri E. Håberg<sup>5</sup>,  
7 George Davey Smith<sup>2,3</sup>, Ted Reichborn-Kjennerud<sup>12,7</sup>, Per Magnus<sup>5</sup>, Alexandra Havdahl<sup>1,7</sup>

8  
9 **Affiliations**

- 10 1. Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway  
11 2. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK  
12 3. Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.  
13 4. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston,  
14 Massachusetts, US  
15 5. Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway.  
16 6. Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, the  
17 Netherlands  
18 7. Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway  
19 8. Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen,  
20 Norway  
21 9. Children and Youth Clinic, Haukeland University Hospital, Bergen, Norway  
22 10. NORMENT, Institute of Clinical Medicine, University of Oslo and Division of Mental Health  
23 and Addiction, Oslo University Hospital, Oslo, Norway  
24 11. School of Psychological Science, University of Bristol, Bristol, UK  
25 12. Institute of Clinical Medicine, University of Oslo, Oslo, Norway

26  
27 **Keywords:** fertility, alcohol, smoking, BMI, caffeine, Mendelian randomisation, MoBa, MBRN.  
28  
29

30 **Abstract (294 of 300 words)**

31 **Introduction:** Current advice to improve fertility includes reducing alcohol and caffeine  
32 consumption, achieving healthy weight-range, and stopping smoking. Advice is informed by  
33 observational evidence, which is often biased by confounding.

34

35 **Methods:** This study uses data from the Norwegian Mother, Father and Child Cohort Study  
36 (MoBa) and the Medical Birth Registry of Norway. First, we analysed associations between  
37 health behaviours prior to pregnancy (alcohol and caffeine consumption, body-mass index (BMI)  
38 and smoking) and multiple indicators of fertility (including number of children, time to  
39 conception, and miscarriage) ( $n=83,128$  women,  $67,555$  men), adjusting for birthyear, education  
40 and attention deficit and hyperactive-impulsive (ADHD) traits. Second, we used individual-level  
41 Mendelian randomisation (MR) to explore possible causal effects of health behaviours on  
42 fertility outcomes ( $n=27,216$  women,  $26,131$  men). Finally, we performed summary-level MR for  
43 available outcomes ( $n=91,462-1,232,091$ ) and conducted multi-variable MR to control for  
44 education and ADHD liability.

45

46 **Results:** In observational analysis, higher BMI and smoking (and to a lesser extent caffeine)  
47 were predominantly associated with reduced fertility outcomes. Unexpectedly, higher alcohol  
48 consumption was associated with predominantly improved fertility outcomes. There was little  
49 evidence from individual-level MR analyses, except smoking and higher BMI were associated  
50 with younger age at first birth in women (mean difference in years, per SD increase in genetic  
51 score; smoking:  $-2.65$  (95%CI:  $-3.57, -1.73$ ); BMI:  $-0.11$  (95%CI:  $-0.16, -0.08$ )) and men  
52 (smoking:  $-2.82$  (95%CI:  $-4.07, -1.58$ ); BMI:  $-0.17$  (95%CI:  $-0.23, -0.11$ )). These results were  
53 replicated in the summary-level MR analysis, however effects attenuated after adjusting for  
54 education and ADHD liability.

55

56 **Conclusions:** Most observational evidence for associations between health behaviours and  
57 fertility was not supported by MR analyses, suggesting possible residual confounding. Evidence  
58 from MR analyses supported an effect of smoking and higher BMI on younger age of first birth,  
59 but multivariable MR suggested this might be explained by underlying liability to ADHD and low  
60 educational attainment.

## 61 **Introduction**

62 Women struggling to conceive are often advised to engage in healthier lifestyle behaviours, for  
63 example, reducing their alcohol and caffeine consumption, achieving a healthy weight-range and  
64 quitting smoking (1–3). Reviews of observational evidence support an association between  
65 these health behaviours and reduced fertility in women (4), with high alcohol consumption and  
66 smoking associated with reduced likelihood of conception (5–9) and high caffeine consumption  
67 and obesity associated with increased risk of miscarriage (10,11). Guidelines and  
68 epidemiological research on fertility often focus on women (2,12,13) however, it is important to  
69 also consider the impact of health behaviours in the partner. Meta-analyses suggest that  
70 smoking and alcohol consumption can reduce semen quality (14), and paternal obesity has been  
71 associated with reduced likelihood of natural conception (15) and an increased time to  
72 conception (16).

73

74 The majority of evidence to date exploring health behaviours and fertility is observational, and  
75 most studies do not adequately control for confounding (17). For example, health behaviours  
76 often co-occur with other health behaviours which might instead affect fertility (e.g. diet,  
77 physical activity, sleep). Furthermore, health behaviours and fertility outcomes (such as age at  
78 first birth and number of children) share common risk factors including low educational  
79 attainment and liability to inattention and hyperactive-impulsive behaviour (traits of attention  
80 deficit hyperactivity disorder; ADHD)(18–20). In addition to confounding, it is necessary to  
81 account for reverse causation because the stress of being unable to conceive might cause couples  
82 to engage in unhealthy behaviours. We extended previous epidemiological studies using  
83 Mendelian randomisation (MR), which can reduce bias from residual confounding and reverse  
84 causation. MR uses genetic variants (single nucleotide polymorphisms; SNPs) to estimate the  
85 causal effect of an exposure on an outcome (21). Causal evidence could improve fertility  
86 guidance, helping couples successfully conceive as well as removing any unnecessary stress or  
87 guilt around unrelated lifestyle factors. The objective of the current study was to explore the  
88 associations between health behaviours and fertility outcomes in both men and women using an  
89 MR approach. We used the Norwegian Mother, Father and Child (MoBa) pregnancy cohort,  
90 containing around 95,000 mothers and 75,000 partners.

91

## 92 **Methods**

### 93 Sample

94 MoBa is a prospective population-based pregnancy cohort study conducted by the Norwegian  
95 Institute of Public Health. Pregnant women were recruited from all over Norway from 1999-  
96 2008 (22,23). Women consented to participation in 41% of the pregnancies. The cohort includes

97 114 500 children, 95 200 mothers and 75 200 partners. The current study was based on version  
98 12 of the quality-assured data files released for research on January 2019. Questionnaires were  
99 completed across multiple time points during pregnancy and after birth. The present study used  
100 measures from the first questionnaires sent to mothers and fathers, received between 13-17  
101 weeks gestation (hereon referred to as the 15-week questionnaire). The questions related to  
102 previous pregnancies, medical history and medication use, occupation, exposures in the  
103 workplace and home, lifestyle habits and mental health. In addition to questionnaire data,  
104 information on maternal characteristics and pregnancy outcomes of the index pregnancies were  
105 available through data linkage to the Medical Birth Registry Norway (MBRN). MBRN also  
106 provided data on age at first birth and total number of children up to 2018, including (but not  
107 limited to) the MoBa index pregnancy. The establishment of MoBa and initial data collection was  
108 based on a license from the Norwegian Data Protection Agency and approval from The Regional  
109 Committees for Medical and Health Research Ethics. The MoBa cohort is now based on  
110 regulations related to the Norwegian Health Registry Act. The current study was approved by  
111 The Regional Committees for Medical and Health Research Ethics (2016/1702).

112

113 After restricting to those with available exposure and outcome data, a maximum of 83,128  
114 women and 67,555 men were included in multivariable regression analysis. After further  
115 restricting to those individuals with quality-controlled genotype data (see details below), a  
116 maximum of 27,216 women and 26,131 men were available for individual-level MR analysis.  
117 More details of participant exclusion are given in Supplementary Figure S1.

118

#### 119 Health behaviours exposure data

120 For all health behaviours, we used data from the 15-week questionnaire in pregnancy. Women  
121 were asked to report their behaviour 3 months prior to the index pregnancy. Fathers were asked  
122 to report their behaviour 6 months prior to the index pregnancy.

123

124 *Alcohol consumption.* Frequency of alcohol consumption was self-reported on a 7-point scale  
125 (never, less than once a month, 1-3 times a month, once per week, 2-3 times per week, 4-5 times  
126 per week, 6-7 times per week). Women and men were also asked about binge drinking  
127 behaviour: “How often did you drink 5 units or more on one occasion?”. Responses were on a 5-  
128 point scale (never, less than once a month, 1-3 times a month, once a week, several times per  
129 week).

130

131 *Caffeine consumption.* Caffeine consumption in women was calculated from self-reported daily  
132 beverage consumption, where one cup contained 125ml. Caffeine (mg) weights per cup were

133 taken from a previous study of caffeine consumption in MoBa (24). We excluded outliers if they  
134 consumed more than 3.5 litres of any particular drink in a day (28 cups) or if their total caffeine  
135 consumption per day was greater than 1000mg (25). Values were log transformed to adjust for  
136 skewness. In men, beverage consumption was instead measured categorically over a typical  
137 week. Responses were on a 5-point scale (seldom/never, 1-6 cups a week, 1 cup a day, 2-3 cups  
138 a day, 4+ cups a day). Unlike the questionnaire administered to women, men were not directed  
139 as to the volume of the cup, therefore we followed previous calculations (24) and assumed that a  
140 cup was 125ml for coffee (apart from espresso where we assumed a standard single is 30ml)  
141 and 250ml for tea or fizzy drink. Again consumption was weighted by caffeine (mg) (24) and  
142 divided by 7 to estimate mg per day.

143

144 *Smoking behaviour.* Smoking initiation was self-reported: “Have you ever smoked?” where yes  
145 was classed as ever smoking and no was classed as never smoking. Smoking heaviness was self-  
146 reported amongst current smokers as the average number of cigarettes smoked per day prior to  
147 pregnancy. We excluded current smokers who reported smoking no cigarettes.

148

149 *Body mass index (BMI).* Height and weight were self-reported. We used pre-pregnancy values for  
150 women and current (15-weeks) values for men. Women were also asked to report their  
151 partner’s height and weight. These reports were highly correlated with partners’ self-report  
152 ( $r=0.98$  for height and  $r=0.95$  for weight). Therefore, we used the woman’s report of their  
153 partner’s height and weight when the partner’s own report was unavailable. We excluded  
154 outliers for women at height  $<117$  or  $>196$  cm and weight  $<38$  or  $>150$  kg and for men at height  
155  $<136$  or  $>220$  cm and weight  $<50$  or  $>200$  kg as done previously in MoBa (16). From these height  
156 and weight measures we calculated BMI as weight (kg)/height (m<sup>2</sup>). We note BMI is not a health  
157 behaviour itself but is a biomarker which can crudely proxy for healthy behaviours.

158

#### 159 Fertility outcome data

160 All fertility outcomes were self-reported in the 15-week questionnaire by the women apart from  
161 age at first birth and total number of children, which were obtained from the Medical Birth  
162 Registry of Norway (MBRN) (version 12, last update November 2018).

163

164 *Age at first birth.* Age at the time of first recorded child’s delivery obtained from MBRN (not  
165 limited to births recorded in MoBa).

166

167 *Number of children.* The total number of children born to women and men up to November 2018  
168 obtained from MBRN (including but not limited to births recorded in MoBa). Totals of more than  
169 6 children were grouped into a 6+ category to adjust for skewness.

170

171 *Time to conception.* Based upon women's self-report from the 15-week questionnaire referring  
172 to the index pregnancy. If women reported planning their pregnancy, they were asked: "For how  
173 many months did you have regular intercourse without contraception before you became  
174 pregnant?". Options were less than one month, 1-2 months or 3+ months. If it took more than 3  
175 months, then they were asked to state the number of months. We combined anyone taking 12 or  
176 more months to conceive into one group to reduce skewness and treated as a continuous  
177 variable. We have used this woman-reported variable as an outcome in both women and men  
178 under the assumption that couples were conceiving together. In our primary analysis, if the  
179 couple were not trying to conceive then they were set to missing. However, given differences  
180 between couples who planned pregnancy and those who did not, we also conducted a sensitivity  
181 analysis where non-planners were included and assigned them the median time to conception  
182 from the planning group (2 months) (Supplementary Note 2). We also conducted a sensitivity  
183 analysis using dichotomised variables as time to conception was not measured continuously  
184 (Supplementary Note 3).

185

186 *Infertility treatment.* Women self-reported in the 15-week questionnaire: "Have you ever been  
187 treated for infertility?". Responses were binary *yes* or *no*. We did not use this variable as an  
188 outcome in the fathers, as this question did not specify whether the infertility treatment was for  
189 the index pregnancy.

190

191 *Miscarriage.* Women self-reported (in the 15-week questionnaire) ever having had a  
192 miscarriage, defined as any of their previous pregnancies ending in spontaneous abortion at or  
193 before the 20<sup>th</sup> week of pregnancy.

194

195 As a secondary outcome, we also explored the impact of health behaviours on frequency of  
196 sexual intercourse (see Supplementary Note 4).

197

#### 198 Genotype data

199 Blood samples were obtained from both parents during pregnancy and from children (umbilical  
200 cord) at birth (26). For the current study we used MoBa Genetics genotype data release 1.0  
201 (<https://github.com/folkhelseinstituttet/mobagen/wiki/MoBaGenetics1.0>). This release has  
202 genotype data available for 99,137 individuals (mostly family trios). Details of the genotyping

203 and QC procedures are available in Supplementary Note 1 (27) and Supplementary Figure S1.  
204 After quality control and relatedness checks, the remaining samples contained 28,929 women  
205 and 27,723 men.

206

#### 207 Genetic score construction

208 For individual-level MR, our genetic instruments were individual-level genetic scores  
209 constructed using PRSice (28) and genome-wide significant variants were possible. SNPs were  
210 clumped to ensure independence ( $r^2 < 0.01$ , clumping window  $< 1000\text{kb}$ ) and weighted by effect  
211 sizes from discovery GWAS detailed below (selected for being the largest samples using  
212 individuals of European ancestry and not containing the MoBa cohort). Prior to analysis, we  
213 checked that each genetic score explained significant variance in the exposure. This is presented  
214 in Supplementary Table S1 along with number of SNPs passing QC in the MoBa cohort.

215

#### 216 GWAS Summary Statistics for health behaviours

217 The following GWAS summary statistics were used to construct genetic scores for individual-  
218 level MR, and individual SNP effects sizes were used in summary-level MR.

219

220 *Alcohol consumption.* We used two genetic instruments for alcohol consumption: 1) alcohol  
221 consumption frequency and 2) binge drinking. Alcohol consumption frequency was measured as  
222 average number of drinks per week aggregated across types of alcoholic beverage. The GWAS  
223 identified 99 conditionally independent genome-wide significant SNPs in a sample of 941,280  
224 individuals, explaining 2.5% of the variance (29). Binge drinking in the UK Biobank was defined  
225 as “How often do you have six or more drinks on one occasion?”, where a drink was defined as a  
226 unit of alcohol. The GWAS was conducted by the Neale Lab (<http://www.nealelab.is/uk-biobank>  
227 – round 2, August 2018), and variance explained in an independent sample was not reported.  
228 After restricting to independent variants, 4 genome-wide significant SNPs remained. Given the  
229 small number of genome-wide significant SNPs, we also used a relaxed p-value threshold of  
230  $p < 5 \times 10^{-6}$ , for which 12 SNPs were available.

231

232 *Caffeine consumption.* Caffeine consumption was measured as number of cups of coffee per day.  
233 The GWAS identified 6 independent genome-wide significant SNPs in a sample of 91,462 coffee  
234 drinkers of European ancestry (30). Genome-wide significant SNPs explained 1.3% of the  
235 variance in coffee consumption. Given the small number of genome-wide significant SNPs, we  
236 also used a relaxed p-value threshold of  $p < 5 \times 10^{-6}$ , for which 39 SNPs were available.

237

238 *Smoking behaviour.* We used two genetic instruments for smoking behaviour: 1) smoking  
239 initiation and 2) smoking heaviness. Smoking initiation was defined as ever v. never regularly  
240 smoking (more than 100 cigarettes ever or having ever been a daily smoker). The GWAS of  
241 smoking initiation identified 378 conditionally independent genome-wide significant SNPs, in a  
242 sample of 1,232,091 individuals, which explained 4% of the variance (29). Smoking heaviness  
243 was defined as average number of cigarettes smoked per day. The GWAS of smoking heaviness  
244 identified 55 conditionally independent genome-wide significant SNPs in a sample of 337,334  
245 ever smokers, which explained 4% of the variance (29). We also conducted a single-SNP analysis  
246 of rs16969968 genotype from the CHRNA5 gene, known to reduce nicotine aversion and  
247 consequently increase cigarettes smoked per day (31) (see Supplementary Note 5).

248

249 *Body mass index (BMI).* The most recent GWAS of adult BMI identified 941 independent genome-  
250 wide significant SNPs in a sample of 681,275 individuals of European ancestry, which explain  
251 6% of the variance in BMI (32).

252

### 253 Statistical analysis

254 All analyses were conducted in R version 4.0.3 (33) and performed separately for women and  
255 men. We corrected for multiple testing using a Bonferroni correction of 0.05/48 tests (6  
256 exposures and 8 outcomes) which resulted in an adjusted p-value of  $p < 0.001$ .

257

258 ***Stage 1. Multivariable regression analyses.*** We first explored the association between each of  
259 the fertility outcomes using correlation for continuous traits, chi-squared tests for binary traits  
260 and independent t-tests when one trait was continuous and the other binary. Second, we  
261 conducted linear regressions for each of the continuous fertility outcomes and logistic  
262 regressions for each of the binary fertility outcomes to see if we observed the expected  
263 association between health behaviours and fertility. Results are presented unadjusted and  
264 adjusted for birth year and educational attainment (at around 15-weeks pregnancy). An  
265 additional adjustment for ADHD traits was included as a sensitivity analysis to proxy for ADHD  
266 liability. ADHD traits were measured when the index child was age 3 years using the Adult ADHD  
267 Self-Report Scale (34).

268

269 ***Stage 2. Individual-level Mendelian randomisation.*** MR can be implemented as an  
270 instrumental variable analysis using genetic variants to proxy for an exposure. It can be used to  
271 estimate a causal effect of the exposure (health behaviours) on the outcome (fertility) providing  
272 certain assumptions are satisfied (21,35). The three core assumptions for valid causal inference  
273 are: 1) relevance – the genetic instrument must be robustly associated with the exposure, 2)



274 independence – there should be no confounding between the genetic instrument and outcome,  
275 and 3) exclusion-restriction – the genetic instrument must only be associated with the outcome  
276 via the exposure. Additionally, for results to generalise to other populations, there must be no  
277 “defiers” – individuals whose exposure is opposite to their genetic predisposition (35). Our  
278 individual-level MR analysis used individual-level genetic scores (with weights from external  
279 independent GWAS) in an instrumental variable regression controlling for age, genotyping chip  
280 and top 10 principal components of population structure. The genetic scores are first regressed  
281 on the exposure, and then predicted values are regressed onto the outcome. Analyses were  
282 performed using the *ivreg* command from the Applied Econometrics with R (*AER*) package for R.  
283 For continuous outcomes, betas are the mean difference per SD increase in the genetic score.  
284 When outcomes are binary, estimates approximate the risk difference per SD increase in genetic  
285 score.

286

287 *Sensitivity analyses.* We checked for evidence of assortative mating by using the woman’s genetic  
288 score to predict the men’s health behaviours and vice versa and by estimating the correlation  
289 between the genetic score of women and men. We checked for evidence of possible pleiotropy  
290 by testing whether each of the genetic scores predicted any known confounders of the exposure-  
291 outcome association (e.g., other health behaviours, income, age) and compared these estimates  
292 with the estimated association between observed exposures on confounders. Pleiotropy occurs  
293 when one genetic variant influences multiple phenotypes. If these other phenotypes are not on  
294 the causal pathway from exposure to outcome (horizontal pleiotropy) then the independence  
295 and exclusion-restriction assumptions could be violated, and genetic variants are invalid. Where  
296 there was evidence for a causal effect in the individual-level MR analysis, we followed up with  
297 additional summary-level MR sensitivity analyses (MR Egger (36), weighted median (37) and  
298 weighted mode (38)) which make different assumptions about the nature of pleiotropy. A  
299 consistent direction of effect across the different MR sensitivity analyses gives us more  
300 confidence that the effects are not due to pleiotropy. We also calculated the MR Egger intercept.  
301 If the intercept is significantly different from zero, then this suggests significant directional  
302 pleiotropy may be biasing the estimate. To conduct these summary-level sensitivity MR  
303 analyses, we generated SNP-outcome association results for the MoBa cohort using the *assoc*  
304 command for plink (version 2) (39).

305

306 **Stage 3. Summary-level Mendelian randomisation.** The second MR method used was  
307 summary-level MR (i.e. two-sample MR) which uses summary statistics from published GWAS  
308 (40). Here we don’t have an effect estimate for each individual but instead an effect size for each  
309 SNP from the discovery GWAS and an effect size for that SNP in an outcome GWAS. The ratio of

310 these two effect sizes can be meta-analysed across multiple SNPs to give us an estimate of the  
311 causal effect. This is our primary analysis known as the inverse-variance weighted estimate. As  
312 sensitivity analyses, again we conducted MR Egger (36), weighted median (37) and weighted  
313 mode (38), which each make orthogonal assumptions about the nature of pleiotropy.

314

315 Independent summary GWAS data was available for three of the outcomes: age at first birth (N=  
316 170,498), number of children (N=333,628) and miscarriage (N= 78,700), self-reported in the UK  
317 Biobank (41). We used outcome GWAS from the UK Biobank only to prevent sample overlap  
318 with our health behaviour exposure GWAS. For age at first birth and miscarriage, we used GWAS  
319 summary statistics from the MRC IEU Open GWAS project (42). The GWAS for age at first birth  
320 used the self-reported question "*How old were you when you had your first child?*" (field 2754)  
321 asked only to women who had previously indicated that they had given birth to at least one  
322 child. The GWAS for miscarriage used the item "*How many spontaneous miscarriages have you*  
323 *had?*" (field 3839) which was only asked to women who had previously indicated that they had  
324 ever had a miscarriage, abortion or stillbirth (field 2774). Finally, the GWAS for number of  
325 children combined items "*How many children have you fathered?*" (field 2405) in men and "*How*  
326 *many children have you given birth to? (Please include live births only)*" (field 2734) in women  
327 (43). Age at first birth and miscarriage GWAS are in SD units and units for number of children  
328 are number of children.

329

330 For each of the exposures, we used the same health behaviour GWAS as for the individual-level  
331 MR, with the exception of BMI. Here we used an earlier GWAS (44) that did not contain the UK  
332 Biobank to avoid sample overlap which can bias estimates (45). For smoking initiation and  
333 alcohol consumption, we used SNP-exposure estimates from summary statistics with the UK  
334 Biobank and 23andMe removed. Smoking heaviness could not be used as an exposure as the  
335 fertility outcome GWASs contain both smokers and non-smokers. Binge drinking could not be  
336 included as this GWAS was also conducted in the UK Biobank.

337

338 *Sensitivity analyses.* The Cochran's Q test of heterogeneity was conducted to estimate possible  
339 pleiotropy and the MR Egger intercept was estimated to test for bias from directional horizontal  
340 pleiotropy. The regression dilution  $I^2_{GX}$  was calculated to assess the suitability of the MR Egger  
341 effect estimate and a SIMEX correction applied where necessary (46). Steiger filtering was  
342 conducted to check for evidence of reverse causation (47). The mean F statistic was calculated as  
343 an indicator of instrument strength, where  $F < 10$  is considered to indicate a weak instrument.  
344 Where there was evidence for a causal effect, we conducted multivariable MR analysis (48) to

345 explore possible pleiotropy via education and ADHD liability (see Supplementary Materials for  
346 details and Table S20).

347

## 348 **Results**

349 Average health behaviours and fertility outcomes were relatively consistent between the full  
350 MoBa sample and the genotyped sub-sample (Table 1). Prevalence of smoking, high alcohol  
351 consumption, high caffeine consumption and high BMI were greater in men than women prior to  
352 pregnancy and men were older on average than women. Fertility outcomes did not differ  
353 between women and men, with the exception of age at first birth which was older in men.  
354 Results for frequency of sexual intercourse, and dichotomised time to conception are given in  
355 Supplementary Notes 3 and 4. Associations between the different fertility outcomes found that a  
356 younger age at first birth was associated with having more children in total, shorter time to  
357 conception, being less likely to miscarry and less likely to have infertility treatment  
358 (Supplementary Table S2). In the following section we highlight results that passed Bonferroni  
359 correction.

360

### 361 Stage 1. Multivariable Regression Associations

362 The results of the observed associations between the health behaviours and fertility outcomes  
363 (adjusted for birth year and education) are given in Figure 1 (Tables S3 and S4).

364

365 *Alcohol Consumption.* Units are per category increase in self-reported alcohol consumption or  
366 binge drinking. In women's adjusted analyses, greater frequency of alcohol consumption and  
367 binge drinking were both associated with having fewer children (alcohol frequency: -0.090  
368 children, 95% CI: -0.096, -0.085; binge drinking: -0.111 children, 95% CI: -0.118, -0.104) and an  
369 older age at first birth (alcohol frequency: 0.492 years, 95% CI: 0.468, 0.516; binge drinking:  
370 0.333 years, 95% CI: 0.304, 0.362). However, those who consumed more alcohol were less likely  
371 to have had infertility treatment (alcohol frequency: OR = 0.867, 95% CI: 0.847, 0.888; binge  
372 drinking: OR = 0.822, 95% CI: 0.798, 0.847). Binge drinking (but not alcohol frequency) was  
373 associated with decreased odds of miscarriage (OR = 0.959, 95% CI: 0.938, 0.981). Neither  
374 alcohol frequency nor binge drinking were associated with time to conception after correction  
375 for multiple testing. The pattern of results was consistent for outcomes available in men.

376

377 *Caffeine consumption.* Units are per unit increase in log transformed mg of caffeine per day. In  
378 men, higher caffeine consumption was associated with a younger age at first birth (-0.100 years,  
379 95% CI: -0.139, -0.061) and being less likely to have fertility treatment (OR = 0.951, 95% CI:

380 0.931, 0.972). Caffeine consumption was not associated with any of the other fertility outcomes  
381 in men nor with any fertility outcomes in women correction for multiple testing.

382

383 *Smoking behaviour.* Ever smoking was associated with having fewer children in women (-0.101  
384 children, 95% CI: -0.114, -0.089), a younger age at first birth in both women and men (women: -  
385 0.155 years, 95% CI: -0.208, -0.103; men: -0.218 years, 95% CI: -0.284, -0.151), and an increased  
386 time to conception in women (0.095 months, 95% CI: 0.038, 0.151). Smoking heaviness was also  
387 associated with having fewer children in women (-0.008 children, 95% CI: -0.010, -0.005) and  
388 longer time to conception (women: 0.029 months, 95% CI: 0.016, 0.042; men: 0.021 months,  
389 95% CI: 0.009, 0.034).

390

391 *Body mass index.* Units are per kg/m<sup>2</sup> increase in BMI. Higher BMI was associated with worse  
392 outcomes across all indicators for fertility in both sexes, including having fewer children  
393 (women: -0.004 children, 95% CI: -0.005, -0.002; men: -0.006 children, 95% CI: -0.008, -0.004)  
394 and taking longer to conceive (women: 0.061 months, 95% CI: 0.055, 0.068; men: 0.045 months,  
395 95% CI: 0.035, 0.054). For women, having a higher BMI was also associated with being more  
396 likely to have fertility treatment (OR: 1.032, 95% CI: 1.026, 1.037) and being more likely to have  
397 a miscarriage (OR: 1.008, 95% CI: 1.004, 1.013). In men, higher BMI was associated with an  
398 older age at first birth (0.026 years, 95% CI: 0.016, 0.035). The only exception to this pattern of  
399 worsened fertility was an association between higher BMI in women and younger age of first  
400 birth (-0.017 years, 95% CI: -0.023, -0.011).

401

402 These results were relatively consistent, but less precise to those observed when restricting the  
403 analysis to the genotyped sample only and after adjustment for ADHD traits (Supplementary  
404 Tables S5 and S6).

405

#### 406 Stage 2. Individual-level Mendelian randomisation

407 Genetic scores were associated with the exposures in MoBa with the exception of binge drinking,  
408 which was therefore not used in further analyses (Table S1). Results of the individual level MR  
409 are presented in Figure 1 (and Tables S7 for women and S8 for men). All genetic scores were  
410 standardised, so that the results are per standard deviation increase in genetic score for the  
411 exposure. Genetic liability for higher BMI was associated with a younger age of first birth in both  
412 women (-0.113 years, 95% CI: -0.153, -0.073) and men (-0.171 years, 95% CI: -0.232, -0.111).  
413 Genetic liability for smoking initiation was also strongly associated with a younger age of first  
414 birth in both women (-2.647 years, 95% CI: -3.565, -1.730) and men (-2.824 years, 95% CI: -  
415 4.072, -1.575). However, stratified by smoking status, rs16969968 genotype was only associated

416 with age at first birth in never smokers, suggesting possible pleiotropy (Supplementary Note 5  
417 and Figure S2). We found no robust evidence for association between any of the other exposure  
418 and outcomes using individual-level MR.

419

420 For both BMI and smoking initiation genetic scores there was evidence of assortative mating  
421 (Table S9) and associations between these genetic scores and the other health behaviours and  
422 household income (Table S10). Additional sensitivity analyses which are more robust to  
423 pleiotropy were consistent and there was no evidence for bias from directional horizontal  
424 pleiotropy (Table S11). Sensitivity analyses in women including non-planners with a median  
425 time to conception suggested an association between higher alcohol consumption and decreased  
426 time to conception (-2.653 months, 95% CI: -4.193, -1.112) and between higher BMI and increased  
427 time to conception (0.969 months, 95% CI: 0.811, 1.129), that were not observed when restricting  
428 to planners (see Supplementary Note 2 for more details).

429

### 430 Stage 3. Summary-level Mendelian randomisation

431 There was strong evidence for an effect of smoking initiation on younger age at first birth in  
432 women (-0.661, 95% CI: -0.757, -0.566), a greater number of children in both men and women  
433 (0.280, 95% CI: 0.205, 0.355) and fewer miscarriages in women (-0.123, 95% CI: -0.182, -0.064)  
434 (see Table 2). All pleiotropy robust sensitivity analyses showed consistently strong evidence  
435 with the same direction of effect. There was evidence of significant heterogeneity (Table S16)  
436 but the MR Egger intercept suggested that these results were not biased by directional  
437 pleiotropy (Table S17). Steiger filtering indicated that the majority of SNPs explained more  
438 variance in smoking initiation than the outcomes (Table S18), suggesting reverse causation is  
439 unlikely.

440

441 There was also some evidence for an effect of higher BMI on younger age at first birth in women  
442 (-0.108, 95% CI: -0.16, -0.056). All MR sensitivity analyses showed a consistent direction of  
443 effect apart from the weighted mode which could indicate possible pleiotropy (Table 2). There  
444 was strong evidence of heterogeneity (Table S16), but the MR Egger intercept did not suggest  
445 this was due to bias from horizontal pleiotropy (Table S17) and there was no evidence of  
446 reverse causation from Steiger filtering (Table S18). There was no evidence for an effect of  
447 alcohol consumption or caffeine consumption on reproductive behaviours. All genetic  
448 instruments had F-statistic > 10 apart from that for smoking initiation (Table S19).

449

450 We conducted multivariable MR analysis (48) to estimate the direct effects of smoking initiation  
451 and BMI on age at first birth in women after accounting for education and ADHD liability (see

452 Supplementary Materials for details and Table S20). Effect estimates for smoking initiation  
453 attenuated after adjustment but there was still evidence for an effect (after adjusting for ADHD: -  
454 0.435, 95% CI: -0.591, -0.279; after adjusting for education: -0.403, 95% CI: -0.527, -0.279). Effect  
455 estimates for BMI were also attenuated, resulting in weak evidence for an effect (ADHD: -0.513,  
456 95% CI: -0.106, 0.003; education: -0.056, 95% CI: -0.113, 0.0008).

457

## 458 **Discussion**

459 This study explored the role of multiple health behaviours on fertility outcomes in men and  
460 women using the MoBa cohort. We extended previous research by including men as well as  
461 women, and strengthening causal inference using MR. We found evidence from multivariable  
462 regression for an association between higher BMI and smoking (and to a lesser extent caffeine)  
463 prior to pregnancy on reduced fertility outcomes. Our MR results were not in line with the  
464 observational associations, finding only support for an effect of higher BMI and smoking  
465 initiation on a younger age at first birth, with evidence of possible horizontal pleiotropy from  
466 ADHD liability and education.

467

468 Frequency of alcohol consumption was associated with being less likely to have infertility  
469 treatment in multivariable regression analyses for both women and men. In contrast, previous  
470 meta-analyses have shown a dose-response relationship between alcohol consumption and  
471 reduced likelihood of conception in women (5) and reduced semen quality in men (14).  
472 However, it is important to note that our analyses were conducted in a pregnancy cohort, so do  
473 not capture those who failed to conceive, introducing possible selection bias. Furthermore,  
474 levels of alcohol consumption were low in our sample, and are possibly below the threshold  
475 which affects fertility (49). To capture more harmful drinking behaviours, we used a measure of  
476 binge drinking, which showed weaker associations, supporting our interpretation. It might be  
477 feasible that low levels of alcohol consumption are not detrimental for fertility, but it seems  
478 unlikely that alcohol could improve fertility, as observed here. Highly confounded multivariable  
479 regression associations of alcohol consumption are a common phenomenon, with low levels of  
480 consumption often being associated with positive outcomes due to confounding by socio-  
481 economic position, or due to never drinkers being a selected group (50). Bias from confounding  
482 is supported by our MR analyses, finding no evidence for effects of alcohol consumption on any  
483 fertility outcomes. Alternatively, our observational results might be affected by reporting bias, as  
484 underreporting of alcohol consumption might be more pronounced amongst those struggling to  
485 conceive. Unfortunately, we were unable to follow up binge drinking associations with MR, as  
486 the genetic instrument did not predict binge drinking in the MoBa cohort. This is not altogether  
487 surprising, as MoBa represents a very different sample to that of the UK Biobank, where the

488 binge drinking GWAS was conducted: 1) the UK Biobank were aged between 40 and 69 years at  
489 recruitment (51) compared with an average age of 31 years in MoBa mothers and 33 years in  
490 MoBa fathers, 2) response rate was higher for the MoBa cohort (41%) compared to the UK  
491 Biobank (5%) (22, 52), 3) drinking behaviours differ greatly between the UK and Norway (53),  
492 with only 6.31% of MoBa mums reporting binge drinking at least weekly in the 3 months prior  
493 to pregnancy, and 4) MoBa is a pregnancy cohort rather than a population cohort, capturing  
494 only fertile couples.

495

496 There was no evidence for an association between caffeine consumption and miscarriage risk, in  
497 contrast to what has been reported in several meta-analyses (54). This could be because our  
498 study explored reported caffeine consumption levels prior to pregnancy rather than during  
499 pregnancy, which has been the primary focus of most previous meta-analyses (54).

500 Alternatively, it could be due to social patterning of caffeine consumption in Norway, with higher  
501 consumption associated with lower levels of education (24) and consequently a younger age at  
502 first birth. Older age is a strong predictor of miscarriage risk (55), so education could be masking  
503 the association. MR results (which are more robust to bias from confounding) did not find  
504 evidence for a causal effect, so it could be possible that previous associations were due to  
505 confounding from other lifestyle factors (56). However, it is important to note that the caffeine  
506 genetic instrument was the weakest, explaining only 0.1-0.2% of the variance. Due to this weak  
507 instrument bias, causal effects cannot be ruled out (57) and replication is warranted.

508

509 We found evidence for an association between smoking heaviness and increased time to  
510 conception in both women and men. This is supported by a previous meta-analysis finding  
511 increased odds of infertility in smokers compared to non-smokers (6). Smoking is hypothesised  
512 to negatively affect sperm production and quality in males (58) and to affect the follicles and  
513 hormone levels in females (49). However, the effect sizes observed in the current study were  
514 quite small and were not supported by the MR results.

515

516 Finally, we saw evidence from multivariable regressions for associations between higher BMI  
517 and reduced fertility, including taking longer to conceive and an increased risk of miscarriage.  
518 These associations are thought to be due to obesity leading to hormone imbalances and  
519 ovulatory dysfunction (similar to very low body weights) (59). However, we did not see  
520 evidence that this effect was causal in the MR analyses. This could be due to confounding, or due  
521 to the MR analysis being unable to detect non-linear effects. Another MR analysis of BMI on  
522 subfertility in the MoBa cohort did find evidence for non-linear effects (60). Intervention  
523 evidence is mixed, with a meta-analysis including randomised control trials finding evidence

524 that interventions for weight loss increased chance of pregnancy but did not affect risk of  
525 miscarriage (61).  
526  
527 A review of studies into health behaviours and fertility concluded that the evidence was robust  
528 for smoking and higher weight reducing fertility (49). However, in the current study, our MR  
529 results did not find evidence to support this effect. In both the individual-level and summary-  
530 level MR we saw consistent evidence for an effect of smoking initiation and higher BMI on  
531 having a younger age at first birth. There is also the possibility that these results are due to  
532 selection bias: both the GWAS for age at first birth and being a participant in the MoBa sample  
533 are conditioned on having had at least one child. If BMI and smoking are associated with reduced  
534 fertility, perhaps only those who had children younger were able to conceive and consequently  
535 be in the sample. However, if this were the case, we might expect to also see effects of smoking  
536 and BMI on other indicators of fertility. An alternative explanation could be that our estimates  
537 are biased by horizontal pleiotropy. We used a range of sensitivity analyses to explore potential  
538 assumption violation by horizontal pleiotropy. Methods which are agnostic to the specific  
539 sources of pleiotropy, (for example, MR-Egger), suggested that the results were not importantly  
540 biased by unbalanced horizontal pleiotropy. However, our exploratory multivariable MR  
541 analyses did show strong attenuation of the effects of BMI and smoking initiation on age at first  
542 birth after accounting for ADHD liability (proxy for ADHD traits of inattentive and hyperactive-  
543 impulsive behaviour) and educational attainment (known risk factors for age at first birth). This  
544 is likely because ADHD traits affect smoking behaviour and some of the smoking genetic variants  
545 relate to smoking via their relationship to ADHD traits. Previous studies have shown a strong  
546 association between the smoking initiation instrument and risk-taking behaviours including  
547 number of sexual partners (62), which could increase the likelihood of having children younger.  
548 Previous Mendelian randomisation studies have also found evidence for bi-directional causal  
549 effects between smoking and education (63,64), smoking and ADHD (65), BMI and education  
550 (66,67) and BMI and ADHD (68). Bi-directional effects between the exposure and the non-  
551 exposure traits can make it difficult to disentangle horizontal from vertical pleiotropy (69).  
552 However, several of these previous studies did find evidence of horizontal pleiotropy, especially  
553 for the smoking initiation instrument (64,65,68). Therefore, we conclude that horizontal  
554 pleiotropy is the most plausible pathway. If there is indeed horizontal pleiotropy from ADHD  
555 liability and education, then direct effects from MVMR accounting for these traits will be closer  
556 to the true causal effect. Finally, there was also strong evidence for assortative mating for both  
557 BMI and smoking initiation instruments which can bias MR results, even for methods which are  
558 robust to horizontal pleiotropy (70).  
559



560 Strengths and Limitations

561 The current study has several strengths. The majority of epidemiological research to date has  
562 focused on women (13), but we also included fathers in our analysis. We observed a similar  
563 pattern of results for women and men in our multivariable regression analyses, despite different  
564 hypothesised mechanisms. This supports our interpretation of widespread confounding,  
565 because pathways through different mechanisms are unlikely to have the same magnitude of  
566 effect (71). Second, this was a very large sample of genotyped individuals with detailed  
567 measures of fertility. Third, we combined multivariable regression and MR methods which each  
568 rely on different assumptions and therefore triangulating across them can strengthen causal  
569 inference. Finally, we included a wide range of health behaviours and different indicators of  
570 fertility.

571

572 This study does have several limitations. First, all MoBa participants were recruited during  
573 pregnancy (12-18 weeks gestation). This means that we are unable to capture the full range of  
574 fertility in the population. Those who never managed to conceive were not observed, and this  
575 could induce selection bias. Second, multivariable regression analyses were cross-sectional and  
576 it is therefore difficult to assess temporality for these associations. Specifically, health  
577 behaviours were retrospectively reported about behaviours 3 or 6 months prior to the index  
578 pregnancy however some couples had been trying to conceive for longer than 6 months.  
579 Furthermore, variables from the Medical Birth Registry of Norway (age at first birth and total  
580 number of children) are across all births, and therefore health behaviours may have differed  
581 compared to before the index pregnancy. Relatedly, there was also a difference between those in  
582 the sample who were planning to conceive compared with those who were not. We hypothesise  
583 that planners are more likely to be cautious about their health behaviours, especially if they have  
584 been having trouble conceiving and have been advised to quit smoking, stop drinking and lose  
585 weight. This could lead to reverse or weakened patterns of association in the multivariable  
586 regression analyses. However, MR would be robust to this type of bias, given that genetic  
587 propensities to health behaviours are fixed at birth. This might explain the different pattern of  
588 results between the multivariable regression and MR analyses. Third, for this paper we have  
589 assumed that partners in MoBa were men. In our genetic analyses this is the case, because  
590 individuals who were not chromosomally XY were removed from the father analysis. However,  
591 in the observational analysis, some partners might have been female partners of the mother, and  
592 these individuals could not be identified. Finally, this study conducted triangulation within a  
593 single sample, and an approach for stronger causal inference would be to triangulate across  
594 different samples in future studies.

595

596 Conclusions

597 For accurate fertility guidance it is extremely important to establish causality. The majority of  
598 associations between health behaviours and fertility observed here, were not replicated using  
599 MR methods. This could suggest that previous observational associations were due to  
600 confounding or other sources of bias, and consequently improving these health behaviours may  
601 not increase the likelihood of conception. We found evidence of potential horizontal pleiotropy,  
602 as our genetic instruments for smoking initiation and BMI were also capturing educational  
603 attainment and ADHD liability. Therefore, triangulation across a broader range of methods,  
604 including those not susceptible to pleiotropy, are required to establish causality.  
605

## 606 **Funding and Acknowledgements**

607 This work was funded by a Gro Harlem Bruntland Scholarship from the Centre of Fertility and  
608 Health NIPH to Dr Wootton. Dr Wootton is currently funded by a postdoctoral fellowship from  
609 the South-Eastern Regional Health Authority (2020024). AH was supported by a career grant  
610 from the South-Eastern Regional Health Authority (2020022), and by the Research Council of  
611 Norway (300668, 288083, 274611). TRK was supported by the Research Council of Norway  
612 (274611). PM and SEH were partly funded by the Research Council of Norway (262700 and  
613 320656). GDS and MRM is part of the Medical Research Council Integrative Epidemiology Unit  
614 (MRC IEU) at the University of Bristol (MM\_UU\_00011/1, MM\_UU\_00011/7). JLT was supported  
615 by a Veni grant from the Netherlands Organization for Scientific Research (NWO: grant number  
616 016.Veni.195.016) and by the Foundation Volksbond Rotterdam. RBL is supported by the  
617 National Institute of Mental Health grant (R01 MH101269).

618

619 This study includes data from the Norwegian Mother, Father and Child Cohort Study (MoBa)  
620 conducted by the Norwegian Institute of Public Health. The Norwegian Mother, Father and Child  
621 Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the  
622 Ministry of Education and Research. We are grateful to all the participating families in Norway  
623 who take part in this on-going cohort study. We thank the Norwegian Institute of Public Health  
624 (NIPH) for generating high-quality genomic data. This research is part of the HARVEST  
625 collaboration, supported by the Research Council of Norway (#229624). For providing genotype  
626 data we thank deCODE Genetics, and the NORMENT Centre for providing genotype data, funded  
627 by the Research Council of Norway (#223273), South East Norway Health Authority and KG  
628 Jebsen Stiftelsen. We further thank the Center for Diabetes Research, the University of Bergen  
629 for providing genotype data and performing quality control and imputation of the data funded  
630 by the ERC AdG project SELECTIONPREDISPOSED, Stiftelsen Kristian Gerhard Jebsen, Trond  
631 Mohn Foundation, the Research Council of Norway, the Novo Nordisk Foundation, the University  
632 of Bergen, and the Western Norway health Authorities (Helse Vest). This research was also  
633 supported by the Research Council or Norway through its Centres of Excellence funding scheme  
634 (project No 262700). MCM has funding from the European Research Council (ERC) under the  
635 European Union's Horizon 2020 research and innovation program (grant agreement No  
636 947684).

637

## 638 **Data availability**

639 The consent and ethical approvals for MoBa does not allow storage of data in repositories or  
640 journals, but it is possible to apply for access to summary statistics datasets for replication or  
641 reproduction of studies by sending an application to [datatilgang@fhi.no](mailto:datatilgang@fhi.no). Data access requires

642 approval from The Regional Committee for Medical and Health Research Ethics in Norway and  
643 an agreement with MoBa.

644

645 **Ethical statement**

646 The establishment and data collection in MoBa was previously based on a license from the  
647 Norwegian Data protection agency and approval from The Regional Committee for Medical  
648 Research Ethics, and it is now based on regulations related to the Norwegian Health Registry  
649 Act. The current study was approved by The Regional Committees for Medical and Health  
650 Research Ethics (2016/217). All GWAS summary statistics are publicly available.

651

652

653

654 **References**

- 655
- 656 1. Kamel RM. Management of the infertile couple: an evidence-based protocol. *Reprod Biol*  
657 *Endocrinol.* 2010 Mar 6;8:21.
  - 658 2. Norwegian Directorate of Health. Gode levevaner før og i svangerskapet [Internet]. 2018.  
659 Available from: [https://www.helseidirektoratet.no/brosjyrer/gode-levevaner-for-og-i-svangerskapet/Gode%20levevaner%20f%C3%B8r%20og%20i%20svangerskapet%20-%20engelsk.pdf/\\_/attachment/inline/73fa64c1-16f3-4052-851c-6a3b2e13ef12:33aa515ffb4a714584b93e4fad70568ec85f4fd0/Gode%20levevaner%20f%C3%B8r%20og%20i%20svangerskapet%20-%20engelsk.pdf](https://www.helseidirektoratet.no/brosjyrer/gode-levevaner-for-og-i-svangerskapet/Gode%20levevaner%20f%C3%B8r%20og%20i%20svangerskapet%20-%20engelsk.pdf/_/attachment/inline/73fa64c1-16f3-4052-851c-6a3b2e13ef12:33aa515ffb4a714584b93e4fad70568ec85f4fd0/Gode%20levevaner%20f%C3%B8r%20og%20i%20svangerskapet%20-%20engelsk.pdf)  
660  
661  
662  
663
  - 664 3. Boivin J, Takefman J, Braverman A. The fertility quality of life (FertiQoL) tool: development  
665 and general psychometric properties. *Hum Reprod Oxf Engl.* 2011 Aug;26(8):2084–91.
  - 666 4. Anderson K, Nisenblat V, Norman R. Lifestyle factors in people seeking infertility  
667 treatment—a review. *Aust N Z J Obstet Gynaecol.* 2010;50(1):8–20.
  - 668 5. Fan D, Liu L, Xia Q, Wang W, Wu S, Tian G, et al. Female alcohol consumption and  
669 fecundability: a systematic review and dose-response meta-analysis. *Sci Rep.*  
670 2017;7(1):13815.
  - 671 6. Augood C, Duckitt K, Templeton AA. Smoking and female infertility: a systematic review  
672 and meta-analysis. *Hum Reprod.* 1998 Jun 1;13(6):1532–9.
  - 673 7. Munafo M, Murphy M, Whiteman D, Hey K. Does cigarette smoking increase time to  
674 conception? *J Biosoc Sci.* 2002;34(1):65–73.
  - 675 8. Cooper AR, Moley KH. Maternal tobacco use and its preimplantation effects on fertility:  
676 more reasons to stop smoking. In: *Seminars in reproductive medicine.* \copyright Thieme  
677 Medical Publishers; 2008. p. 204–12.
  - 678 9. Wesselink AK, Hatch EE, Rothman KJ, Mikkelsen EM, Aschengrau A, Wise LA. Prospective  
679 study of cigarette smoking and fecundability. *Hum Reprod Oxf Engl.* 2019 Mar  
680 1;34(3):558–67.
  - 681 10. Cavalcante MB, Sarno M, Peixoto AB, Araujo Júnior E, Barini R. Obesity and recurrent  
682 miscarriage: A systematic review and meta-analysis. *J Obstet Gynaecol Res.* 2019;45(1):30–  
683 8.
  - 684 11. Li J, Zhao H, Song J-M, Zhang J, Tang Y-L, Xin C-M. A meta-analysis of risk of pregnancy loss  
685 and caffeine and coffee consumption during pregnancy. *Int J Gynecol Obstet.* 2015 Aug  
686 1;130(2):116–22.
  - 687 12. Easey K, Sharp GC. The impact of paternal alcohol, tobacco, caffeine use and physical  
688 activity on offspring mental health: A systematic review and meta-analysis. *medRxiv.* 2021;
  - 689 13. Sharp GC, Lawlor DA, Richardson SS. It’s the mother!: How assumptions about the causal  
690 primacy of maternal effects influence research on the developmental origins of health and  
691 disease. *Soc Sci Med.* 2018;213:20–7.
  - 692 14. Li Y, Lin H, Li Y, Cao J. Association between socio-psycho-behavioral factors and male  
693 semen quality: systematic review and meta-analyses. *Fertil Steril.* 2011;95(1):116–23.

- 694 15. Campbell JM, Lane M, Owens JA, Bakos HW. Paternal obesity negatively affects male  
695 fertility and assisted reproduction outcomes: a systematic review and meta-analysis.  
696 *Reprod Biomed Online*. 2015 Nov 1;31(5):593–604.
- 697 16. Nguyen RH, Wilcox AJ, Skjærven R, Baird DD. Men’s body mass index and infertility. *Hum*  
698 *Reprod*. 2007;22(9):2488–93.
- 699 17. Peck JD, Leviton A, Cowan LD. A review of the epidemiologic evidence concerning the  
700 reproductive health effects of caffeine consumption: A 2000–2009 update. *Food Chem*  
701 *Toxicol*. 2010 Oct 1;48(10):2549–76.
- 702 18. Groenman AP, Janssen TWP, Oosterlaan J. Childhood Psychiatric Disorders as Risk Factor  
703 for Subsequent Substance Abuse: A Meta-Analysis. *J Am Acad Child Adolesc Psychiatry*.  
704 2017 Jul 1;56(7):556–69.
- 705 19. Kong A, Frigge ML, Thorleifsson G, Stefansson H, Young AI, Zink F, et al. Selection against  
706 variants in the genome associated with educational attainment. *Proc Natl Acad Sci*.  
707 2017;114(5):E727–32.
- 708 20. Hvolgaard Mikkelsen S, Olsen J, Bech BH, Obel C. Parental age and attention-  
709 deficit/hyperactivity disorder (ADHD). *Int J Epidemiol*. 2017 Apr 1;46(2):409–20.
- 710 21. Davey Smith G, Ebrahim S. “Mendelian randomization”: can genetic epidemiology  
711 contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003  
712 Feb;32(1):1–22.
- 713 22. Magnus P, Irgens LM, Haug K, Nystad W, Skjærven R, Stoltenberg C. Cohort profile: the  
714 Norwegian mother and child cohort study (MoBa). *Int J Epidemiol*. 2006;35(5):1146–50.
- 715 23. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, et al. Cohort profile update:  
716 the Norwegian mother and child cohort study (MoBa). *Int J Epidemiol*. 2016;45(2):382–8.
- 717 24. Papadopoulou E, Botton J, Brantsæter A-L, Haugen M, Alexander J, Meltzer HM, et al.  
718 Maternal caffeine intake during pregnancy and childhood growth and overweight: results  
719 from a large Norwegian prospective observational cohort study. *BMJ Open*.  
720 2018;8(3):e018895.
- 721 25. Treur JL, Taylor AE, Ware JJ, McMahon G, Hottenga J-J, Baselmans BM, et al. Associations  
722 between smoking and caffeine consumption in two European cohorts. *Addiction*.  
723 2016;111(6):1059–68.
- 724 26. Paltiel L, Anita H, Skjærden T, Harbak K, Bækken S, Kristin SN, et al. The biobank of the  
725 Norwegian Mother and Child Cohort Study – present status. *Nor Epidemiol [Internet]*. 2014  
726 Dec 22 [cited 2021 Jun 9];24(1–2). Available from:  
727 <https://www.ntnu.no/ojs/index.php/norepid/article/view/1755>
- 728 27. Helgeland Ø, Vaudel M, Juliusson PB, Lingaas Holmen O, Juodakis J, Bacelis J, et al. Genome-  
729 wide association study reveals dynamic role of genetic variation in infant and early  
730 childhood growth. *Nat Commun*. 2019 Oct 1;10(1):4448.
- 731 28. Euesden J, Lewis CM, O’Reilly PF. PRSice: Polygenic Risk Score software. *Bioinforma Oxf*  
732 *Engl*. 2015 May 1;31(9):1466–8.

- 733 29. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2  
734 million individuals yield new insights into the genetic etiology of tobacco and alcohol use.  
735 *Nat Genet.* 2019;51:237–44.
- 736 30. Cornelis MC, Byrne EM, Esko T, Nalls MA, Ganna A, Paynter N, et al. Genome-wide meta-  
737 analysis identifies six novel loci associated with habitual coffee consumption. *Mol*  
738 *Psychiatry.* 2015 May;20(5):647–56.
- 739 31. Jensen KP, DeVito EE, Herman AI, Valentine GW, Gelernter J, Sofuoglu M. A CHRNA5  
740 Smoking Risk Variant Decreases the Aversive Effects of Nicotine in Humans.  
741 *Neuropsychopharmacology.* 2015 Nov;40(12):2813–21.
- 742 32. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of  
743 genome-wide association studies for height and body mass index in ~700,000 individuals  
744 of European ancestry. *bioRxiv.* 2018 Mar 2;274654.
- 745 33. R. Core Team. R: A language and environment for statistical computing. R Foundation for  
746 Statistical Computing, Vienna, Austria. 2013. ISBN 3-900051-07-0; 2014.
- 747 34. Kessler RC, Adler L, Ames M, Demler O, Faraone S, Hiripi E, et al. The World Health  
748 Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the  
749 general population. *Psychol Med.* 2005 Feb;35(2):245–56.
- 750 35. Lousdal ML. An introduction to instrumental variable assumptions, validation and  
751 estimation. *Emerg Themes Epidemiol.* 2018 Jan 22;15:1.
- 752 36. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments:  
753 effect estimation and bias detection through Egger regression. *Int J Epidemiol.*  
754 2015;44(2):512–25.
- 755 37. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian  
756 randomization with some invalid instruments using a weighted median estimator. *Genet*  
757 *Epidemiol.* 2016;40(4):304–14.
- 758 38. Hartwig FP, Smith GD, Bowden J. Robust inference in two-sample Mendelian randomisation  
759 via the zero modal pleiotropy assumption. *bioRxiv.* 2017;126102.
- 760 39. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set  
761 for whole-genome association and population-based linkage analyses. *Am J Hum Genet.*  
762 2007;81(3):559–75.
- 763 40. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and  
764 challenges. *Int J Epidemiol.* 2016;45(3):908.
- 765 41. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access  
766 resource for identifying the causes of a wide range of complex diseases of middle and old  
767 age. *PLoS Med.* 2015;12(3):e1001779.
- 768 42. Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, et al. The MRC IEU  
769 OpenGWAS data infrastructure. *bioRxiv.* 2020 Aug 10;2020.08.10.244293.
- 770 43. Lawn RB, Sallis HM, Taylor AE, Wootton RE, Smith GD, Davies NM, et al. Schizophrenia risk  
771 and reproductive success: A Mendelian randomization study. *R Soc Open Sci.* 2018 Jun  
772 28;357673.

- 773 44. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body  
774 mass index yield new insights for obesity biology. *Nature*. 2015;518(7538):197–206.
- 775 45. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample  
776 Mendelian randomization. *Genet Epidemiol*. 2016 Nov;40(7):597–608.
- 777 46. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing  
778 the suitability of summary data for two-sample Mendelian randomization analyses using  
779 MR-Egger regression: the role of the I<sup>2</sup> statistic. *Int J Epidemiol*. 2016;45(6):1961–74.
- 780 47. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely  
781 measured traits using GWAS summary data. Li J, editor. *PLoS Genet*. 2017 Nov  
782 17;13(11):e1007081.
- 783 48. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable  
784 Mendelian randomization in the single-sample and two-sample summary data settings. *Int*  
785 *J Epidemiol*. 2019;48(3):713–27.
- 786 49. Homan GF, Davies M, Norman R. The impact of lifestyle factors on reproductive  
787 performance in the general population and those undergoing infertility treatment: a  
788 review. *Hum Reprod Update*. 2007 May 1;13(3):209–23.
- 789 50. Connor J. The Life and Times of the J-Shaped Curve. *Alcohol Alcohol*. 2006 Nov  
790 1;41(6):583–4.
- 791 51. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource  
792 with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203–9.
- 793 52. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection  
794 bias can substantially influence observed associations. *Int J Epidemiol*. 2018 Feb  
795 1;47(1):226–35.
- 796 53. World Health Organization. Global status report on alcohol and health 2018 [Internet].  
797 Geneva; 2018 [cited 2019 Jan 4]. Available from:  
798 [http://apps.who.int/iris/bitstream/handle/10665/274603/9789241565639-](http://apps.who.int/iris/bitstream/handle/10665/274603/9789241565639-eng.pdf?ua=1)  
799 [eng.pdf?ua=1](http://apps.who.int/iris/bitstream/handle/10665/274603/9789241565639-eng.pdf?ua=1).
- 800 54. James JE. Maternal caffeine consumption and pregnancy outcomes: a narrative review with  
801 implications for advice to mothers and mothers-to-be. *BMJ Evid-Based Med*. 2021 Jun  
802 1;26(3):114–5.
- 803 55. Magnus MC, Wilcox AJ, Morken N-H, Weinberg CR, Håberg SE. Role of maternal age and  
804 pregnancy history in risk of miscarriage: prospective register based study. *BMJ*. 2019 Mar  
805 20;364:l869.
- 806 56. Chen L, Bell EM, Browne ML, Druschel CM, Romitti PA, National Birth Defects Prevention  
807 Study. Exploring maternal patterns of dietary caffeine consumption before conception and  
808 during pregnancy. *Matern Child Health J*. 2014 Dec;18(10):2446–55.
- 809 57. Cornelis MC, Munafo MR. Mendelian Randomization Studies of Coffee and Caffeine  
810 Consumption. *Nutrients*. 2018 Oct;10(10):1343.
- 811 58. Künzle R, Mueller MD, Hänggi W, Birkhäuser MH, Drescher H, Bersinger NA. Semen quality  
812 of male smokers and nonsmokers in infertile couples. *Fertil Steril*. 2003 Feb 1;79(2):287–  
813 91.



- 814 59. Silvestris E, de Pergola G, Rosania R, Loverro G. Obesity as disruptor of the female fertility.  
815 *Reprod Biol Endocrinol*. 2018 Mar 9;16(1):22.
- 816 60. Hernáez Á, Rogne T, Skåra KH, Håberg SE, Page CM, Fraser A, et al. Body mass index and  
817 subfertility: multivariable regression and Mendelian randomization analyses in the  
818 Norwegian Mother, Father and Child Cohort Study. *Hum Reprod*. 2021 Dec 1;36(12):3141–  
819 51.
- 820 61. Best D, Avenell A, Bhattacharya S. How effective are weight-loss interventions for  
821 improving fertility in women and men who are overweight or obese? A systematic review  
822 and meta-analysis of the evidence. *Hum Reprod Update*. 2017 Nov 1;23(6):681–705.
- 823 62. Khouja JN, Wootton RE, Taylor AE, Smith GD, Munafò MR. Association of genetic liability to  
824 smoking initiation with e-cigarette use in young adults: A cohort study. *PLOS Med*. 2021  
825 Mar 18;18(3):e1003555.
- 826 63. Gage SH, Bowden J, Davey Smith G, Munafò MR. Investigating causality in associations  
827 between education and smoking: a two-sample Mendelian randomization study. *Int J*  
828 *Epidemiol*. 2018 Aug 1;47(4):1131–40.
- 829 64. Gage SH, Sallis HM, Lassi G, Wootton RE, Mokrysz C, Davey Smith G, et al. Does smoking  
830 cause lower educational attainment and general cognitive ability? Triangulation of causal  
831 evidence using multiple study designs. *Psychol Med*. 2020 Oct 7;1–9.
- 832 65. Treur JL, Demontis D, Smith GD, Sallis H, Richardson TG, Wiers RW, et al. Investigating  
833 causality between liability to ADHD and substance use, and liability to substance use and  
834 ADHD risk, using Mendelian randomization. *Addict Biol*. 2021 Jan;26(1):e12849.
- 835 66. Böckerman P, Viinikainen J, Pulkki-Råback L, Hakulinen C, Pitkänen N, Lehtimäki T, et al.  
836 Does higher education protect against obesity? Evidence using Mendelian randomization.  
837 *Prev Med*. 2017 Aug;101:195–8.
- 838 67. Hughes A, Wade KH, Dickson M, Rice F, Davies A, Davies NM, et al. Common health  
839 conditions in childhood and adolescence, school absence, and educational attainment:  
840 Mendelian randomization study. *Npj Sci Learn*. 2021 Jan 4;6(1):1–9.
- 841 68. Liu C-Y, Schoeler T, Davies NM, Peyre H, Lim K-X, Barker ED, et al. Are there causal  
842 relationships between attention-deficit/hyperactivity disorder and body mass index?  
843 Evidence from multiple genetically informed designs. *Int J Epidemiol*. 2021 May  
844 17;50(2):496–509.
- 845 69. Yang Q, Sanderson E, Tilling K, Borges MC, Lawlor DA. Exploring and mitigating potential  
846 bias when genetic instrumental variables are associated with multiple non-exposure traits  
847 in Mendelian randomization [Internet]. medRxiv; 2019 [cited 2022 Feb 17]. p. 19009605.  
848 Available from: <https://www.medrxiv.org/content/10.1101/19009605v1>
- 849 70. Hartwig FP, Davies NM, Davey Smith G. Bias in Mendelian randomization due to assortative  
850 mating. *Genet Epidemiol*. 2018 Oct;42(7):608–20.
- 851 71. Smith GD. Assessing Intrauterine Influences on Offspring Health Outcomes: Can  
852 Epidemiological Studies Yield Robust Findings? *Basic Clin Pharmacol Toxicol*.  
853 2008;102(2):245–56.

854  
855



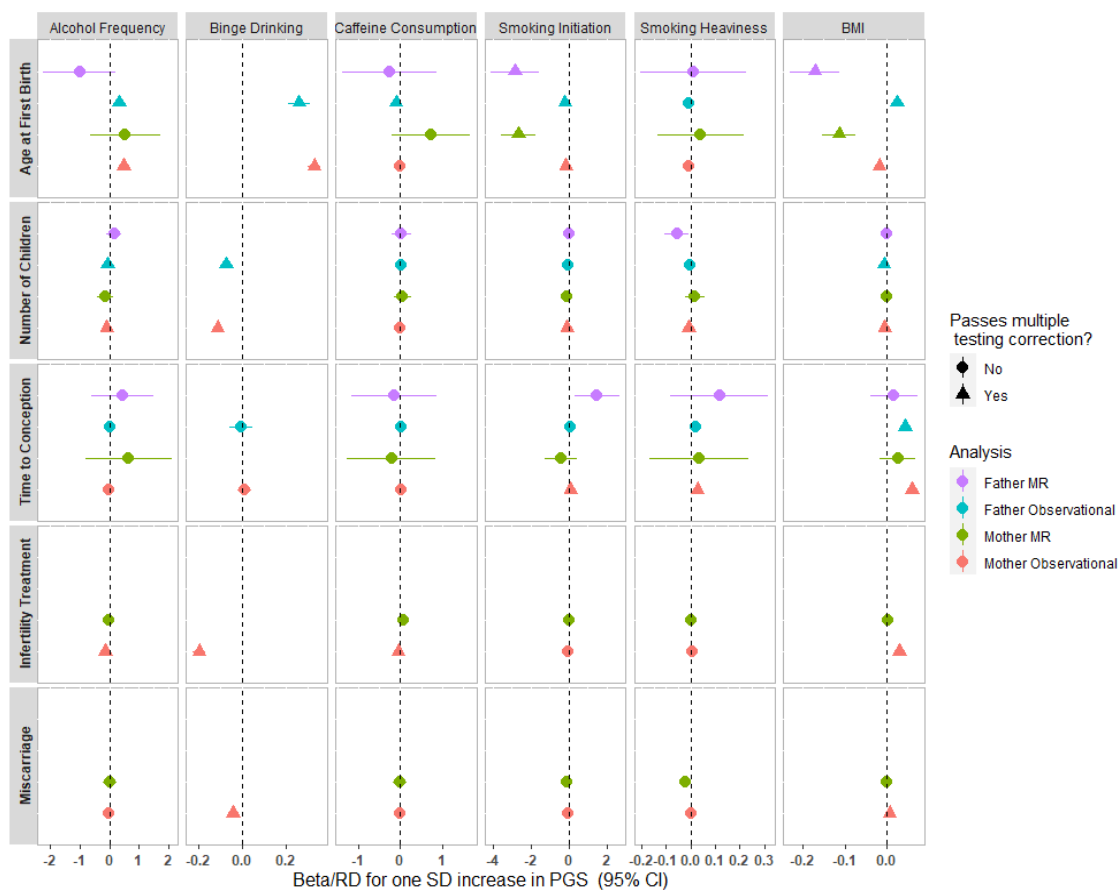
857  
858

**Table 1. Descriptive statistics comparing exposure and outcome data across mothers and fathers.**

	<b>Mothers</b>				<b>Fathers</b>				
	Full Sample		Genotyped Sample		Full Sample		Genotyped Sample		
	<b>N</b>	<b>Mean (SD)/%</b>	<b>N</b>	<b>Mean (SD)/%</b>	<b>N</b>	<b>Mean (SD)/%</b>	<b>N</b>	<b>Mean (SD)/%</b>	
<b>Age (years)</b>	94,418	30.61 (20.66)	28,849	30.34 (15.55)	74,409	33.05 (18.73)	27,598	32.90 (17.44)	
<b>BMI (kg/m<sup>2</sup>)</b>	83,191	24.05 (4.30)	27,272	24.08 (4.25)	67,679*	25.88 (3.34)	26,161*	25.93 (3.33)	
<b>Alcohol Consumption</b>	<i>Total</i>	-	26,819	-	64,023	-	24,775	-	
	<i>Never</i>	5,690	7.15%	1,885	7.03%	1,441	2.25%	473	1.91%
	<i>Less than once a month</i>	23,618	29.69%	7,848	29.26%	11,203	17.50%	4,220	17.03%
	<i>1-3 times a month</i>	28,575	35.92%	9,804	36.56%	23,497	36.70%	9,304	37.55%
	<i>Once per week</i>	14,263	17.93%	4,811	17.94%	15,869	24.79%	6,250	25.23%
	<i>2-3 times per week</i>	6,527	8.20%	2,188	8.16%	10,168	15.88%	3,896	15.73%
	<i>4-5 times per week</i>	724	0.91%	221	0.82%	1,488	2.32%	521	2.10%
	<i>6-7 times per week</i>	165	0.21%	62	0.23%	357	0.56%	111	0.45%
<b>Binge Drinking</b>	<i>Total</i>	78,615	-	26,546	-	28,699	-	10,761	-
	<i>Never</i>	26,229	33.36%	8,707	32.80%	3,001	10.46%	1,058	9.83%
	<i>Less than once a month</i>	31,002	39.44%	10,647	40.11%	11,726	40.86%	4,575	42.51%
	<i>1-3 times a month</i>	16,428	20.90%	5,606	21.12%	9,534	33.22%	3,545	32.94%
	<i>Once a week</i>	4,313	5.49%	1,411	5.32%	3,732	13.00%	1,348	12.53%
	<i>Several times per week</i>	643	0.82%	175	0.66%	706	2.46%	235	2.18%
<b>Smoking Initiation</b>	<i>Total</i>	82,824	-	27,601	-	65,670	-	25,324	-
	<i>Never smokers</i>	41,326	49.90%	13,897	50.35%	30,993	47.20%	12,124	47.88%
	<i>Ever smokers</i>	41,498	50.10%	13,704	49.65%	34,677	52.80%	13,200	52.12%
<b>Smoking Heaviness (cigarettes per day)</b>		16,074	11.41 (5.94)	4,961	11.38 (5.85)	12,606	13.49 (6.39)	4,721	13.42 (6.17)
<b>Caffeine Consumption (mg per day)</b>		75,172	141.60 (138.36)	24,512	139.40 (136.63)	27,501	125.40 (85.60)	10,326	125.79 (84.99)
<b>Age at First Birth (years)</b>		94,740	27.15 (4.65)	28,871	27.32 (4.42)	74,681	29.61 (4.99)	27,617	29.63 (4.82)
<b>Number of children (N children)</b>		94,740	2.55 (0.95)	28,871	2.56 (0.89)	74,681	2.51 (0.91)	27,617	2.52 (0.88)
<b>Time to Conception (months)</b>		63,068	4.90 (7.94)	21,749	4.63 (7.39)	53,516∇	4.86 (7.84)	21,146∇	4.64 (7.40)
<b>Infertility treatment</b>	<i>Total</i>	83,876	-	27,616	-	-	-	-	-
	<i>Never</i>	76,195	90.84%	25,253	91.44%	-	-	-	-
	<i>Ever</i>	7,681	9.16%	2,363	8.56%	-	-	-	-
<b>Miscarriage</b>	<i>Total</i>	55,696	-	17,896	-	-	-	-	-
	<i>Never</i>	39,114	70.23%	12,617	70.50%	-	-	-	-
	<i>Ever</i>	16,582	29.77%	5,279	29.50%	-	-	-	-

859 Note. \* = supplemented with mother's report when father's report was unavailable. ∇ = Father variable reported by the mother.

860



861

862 *Figure 1. The association between health behaviours and fertility outcomes comparing*  
 863 *mothers and fathers, and comparing multivariable regression associations with estimates*  
 864 *of causal effects from Mendelian randomisation. For continuous outcomes, units are betas.*  
 865 *For binary outcomes, regression units are on the log odds scale and MR units are risk*  
 866 *differences.*

867

868

869

870 **Table 2. Summary level Mendelian randomisation results for health behaviours on age at first birth, number**  
 871 **of children and number of miscarriages**  
 872

Exposure	Outcome	Method	N SNP	Beta (95% CI)	P
<b>Drinks per week</b>	Age at First Birth	IVW	94	-0.055 (-0.145, 0.034)	0.22
		MR Egger	94	0.04 (-0.095, 0.175)	0.56
		Weighted median	94	-0.048 (-0.139, 0.042)	0.30
		Weighted mode	94	-0.046 (-0.129, 0.038)	0.29
	Number of children	IVW	83	0.027 (-0.057, 0.110)	0.53
		MR Egger (SIMEX)	83	0.014 (-0.029, 0.058)	0.52
		Weighted median	83	-0.010 (-0.094, 0.074)	0.82
		Weighted mode	83	0.009 (-0.068, 0.086)	0.82
Number of miscarriages	IVW	94	-0.029 (-0.097, 0.038)	0.39	
	MR Egger	94	-0.018 (-0.121, 0.084)	0.73	
	Weighted median	94	0.001 (-0.099, 0.102)	0.98	
	Weighted mode	94	0.01 (-0.084, 0.103)	0.84	
<b>Caffeine</b>	Age at First Birth	IVW	6	-0.036 (-0.13, 0.057)	0.45
		Weighted median	6	0.013 (-0.073, 0.099)	0.77
		Weighted mode	6	0.033 (-0.067, 0.134)	0.54
	Number of children	IVW	6	0.021 (-0.031, 0.073)	0.43
		Weighted median	6	0.021 (-0.047, 0.089)	0.54
	Number of miscarriages	IVW	6	-0.042 (-0.106, 0.022)	0.20
		Weighted median	6	-0.052 (-0.128, 0.025)	0.19
		Weighted mode	6	-0.076 (-0.178, 0.025)	0.20
	<b>Smoking Initiation</b>	Age at First Birth	IVW	322	-0.661 (-0.757, -0.566)
Weighted median			322	-0.504 (-0.603, -0.405)	2.71 x 10 <sup>-23</sup>
Weighted mode			322	-0.555 (-0.785, -0.326)	3.30 x 10 <sup>-6</sup>
Number of children		IVW	323	0.280 (0.205, 0.355)	3.17 x 10 <sup>-13</sup>
		Weighted median	323	0.207 (0.125, 0.289)	7.11 x 10 <sup>-07</sup>
		Weighted mode	323	0.177 (-0.009, 0.363)	0.06
Number of miscarriages		IVW	322	-0.123 (-0.182, -0.064)	4.68 x 10 <sup>-5</sup>
		Weighted median	322	-0.106 (-0.196, -0.017)	0.019
		Weighted mode	322	-0.176 (-0.399, 0.047)	0.124
<b>BMI</b>	Age at First Birth	IVW	95	-0.108 (-0.16, -0.056)	4.32 x 10 <sup>-05</sup>
		MR Egger	95	-0.076 (-0.225, 0.072)	0.32
		Weighted median	95	-0.043 (-0.101, 0.015)	0.14
		Weighted mode	95	0.006 (-0.114, 0.126)	0.92
	Number of children	IVW	91	-0.014 (-0.054, 0.027)	0.51
		MR Egger	91	-0.002 (-0.117, 0.113)	0.98
		Weighted median	91	-0.008 (-0.054, 0.038)	0.73
		Weighted mode	91	-0.013 (-0.094, 0.069)	0.76
	Number of miscarriages	IVW	95	0.004 (-0.033, 0.041)	0.83
		MR Egger	95	0.021 (-0.086, 0.128)	0.70
		Weighted median	95	0.019 (-0.033, 0.072)	0.47
		Weighted mode	95	0.013 (-0.091, 0.116)	0.81

873  
 874 Note. Caffeine genetic instrument using a relaxed p-value threshold of  $p < 5 \times 10^{-6}$ . IVW = inverse-variance  
 875 weighted; BMI = body mass index. The I2GX suggested that the smoking initiation and caffeine genetic  
 876 instruments were not suitable for conducting MR Egger (Supplementary Table S14). Unweighted SIMEX  
 877 corrections were conducted for smoking initiation.