





Mendelian Randomization

Proxy gene-by-environment Mendelian randomization study confirms a causal effect of maternal smoking on offspring birthweight, but little evidence of long-term influences on offspring health

Qian Yang ^{1,2*}, Louise A C Millard^{1,2,3†} and George Davey Smith ^{1,2†}

¹Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK, ²Population Health Sciences, University of Bristol, Bristol, UK and ³Intelligent Systems Laboratory, Department of Computer Science, University of Bristol, Bristol, UK

*Corresponding author. OF 28, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK. E-mail: qian.yang@bristol.ac.uk

†These authors contributed equally.

Editorial decision 4 November 2019; Accepted 18 November 2019

Abstract

Background: A lack of genetic data across generations makes transgenerational Mendelian randomization (MR) difficult. We used UK Biobank and a novel proxy gene-by-environment MR to investigate effects of maternal smoking heaviness in pregnancy on offspring health, using participants' (generation one: G1) genotype (rs16969968 in *CHRNA5*) as a proxy for their mothers' (G0) genotype.

Methods: We validated this approach by replicating an established effect of maternal smoking heaviness on offspring birthweight. Then we applied this approach to explore effects of maternal (G0) smoking heaviness on offspring (G1) later life outcomes and on birthweight of G1 women's children (G2).

Results: Each additional smoking-increasing allele in offspring (G1) was associated with a 0.018 [95% confidence interval (CI): -0.026, -0.009] kg lower G1 birthweight in maternal (G0) smoking stratum, but no meaningful effect (-0.002 kg; 95% CI: -0.008, 0.003) in maternal non-smoking stratum (interaction *P*-value = 0.004). The differences in associations of rs16969968 with grandchild's (G2) birthweight between grandmothers (G0) who did, versus did not, smoke were heterogeneous (interaction *P*-value = 0.042) among mothers (G1) who did (-0.020 kg/allele; 95% CI: -0.044, 0.003), versus did not (0.007 kg/allele; 95% CI: -0.005, 0.020), smoke in pregnancy.

Conclusions: Our study demonstrated how offspring genotype can be used to proxy for the mother's genotype in gene-by-environment MR. We confirmed the causal effect of maternal (G0) smoking on offspring (G1) birthweight, but found little evidence of an effect on G1 longer-term health outcomes. For grandchild's (G2) birthweight, the effect of

grandmother's (G0) smoking heaviness in pregnancy may be modulated by maternal (G1) smoking status in pregnancy.

Key words: Gene \times environment, Mendelian randomization, proxy, maternal smoking, pregnancy

Key Messages

- Our study presents a novel proxy gene-by-environment (G \times E) Mendelian randomization (MR) approach to explore maternal effects on offspring phenotypes when maternal genetic information is unavailable.
- We demonstrated the proxy G \times E MR approach in UK Biobank, showing that heavier maternal smoking led to lower offspring birthweight, which has been demonstrated using several study designs.
- Using this method, we found little evidence that maternal smoking heaviness in pregnancy was associated with offspring's height, body mass index, lung function, risk of asthma, blood pressure, age at menarche, years of education, intelligence score, depression/anxiety or happiness in adulthood.
- Our study also suggested that the effect of grandmother's smoking heaviness in pregnancy on grandchild's birthweight may be modulated by maternal smoking status in pregnancy.
- Further G \times E MR studies with larger sample sizes are needed to assess transgenerational effects with greater statistical power.

Introduction

The developmental origins of health and disease hypothesis proposes that early life experiences, including those *in utero*, can have long-term health effects, and maternal pregnancy exposures are important to long-term health of offspring.¹ Heavier maternal smoking in pregnancy is known to be causally associated with lower offspring birthweight,^{2–6} but its other effects in offspring are less clear. Multivariable regression in observational data showed that heavier maternal smoking during pregnancy was associated with offspring being shorter⁷ and more overweight/obese,^{8,9} and having higher blood pressure,¹⁰ but had mixed associations with age at menarche¹¹ and respiratory,¹² cognitive¹³ and mental health.¹⁴ Heavier maternal smoking in pregnancy has also been associated with higher grandchild's birthweight in certain subpopulations.^{15–17} It is unclear whether these associations reflect a causal effect of maternal smoking in pregnancy, as they may be due to residual confounding. Some studies have assessed this using paternal smoking as a 'negative control', since an effect via uterine environment would be observed in mothers but not fathers, such that similar-magnitude associations would indicate confounding via shared familial, social, environmental and genetic factors.^{2,5,18} Negative control studies suggest little evidence of a causal effect on offspring body mass index (BMI),^{2,5,8} blood pressure^{19,20} and depression.²¹

Mendelian randomization (MR) provides an alternative way to explore this question by using single nucleotide polymorphisms (SNPs) as instrumental variables (IVs) for

an exposure of interest. MR is less prone to confounding, as germline genetic variants are randomly allocated at meiosis and are not influenced by subsequent socioeconomic and health behaviours.^{22,23} MR has been applied in a gene-by-environment (G \times E) framework,^{24,25} which requires variation in the strength of the gene-exposure association across strata of another factor. If there is a causal effect of the IV on the outcome via the exposure of interest, then we would expect the association of the IV with the outcome to vary in proportion to the gene-exposure association. The rs1051730/rs16969968 (*CHRNA5*) SNPs, previously robustly associated with smoking heaviness among smokers,²⁶ have been widely used as IVs for smoking heaviness in G \times E MR studies.^{3,27–29} A causal effect of the smoking heaviness IV on an outcome should be seen among ever but not among never smokers if the effect is via smoking heaviness rather than other pathways.^{24,25} G \times E MR has also been used to assess cross-generational causal effects. A smoking heaviness IV has been associated with lower offspring birthweight among mothers who smoked in pregnancy but not among mothers who did not smoke in pregnancy, suggesting that the genetic instrument affects birthweight through maternal smoking.³

It is usually difficult to investigate transgenerational associations, due to a lack of data across the generations of interest. Thus, previous work has sought to test transgenerational associations using available traits as proxies for unmeasured traits of interest. A Norwegian cohort aimed to examine whether women's smoking in adulthood was related to their mothers' smoking habits (that were not

recorded) and hence used maternal smoking-related mortality as a proxy.³⁰ Recently, a case-control by proxy approach has been proposed.³¹ Participants' genotypes were used to proxy for unavailable parental genotypes, and their associations were tested against parental diagnosis of Alzheimer's disease in UK Biobank,³¹ since Alzheimer's disease was much more prevalent in the parents than in the participants (aged between 40 and 69 at baseline in 2006–10³²). Our study aimed to demonstrate how an analogous approach can be used within a $G \times E$ MR framework to test maternal-offspring effects when maternal genotype is not available, using offspring genotype as a proxy for the maternal genotype. First, we performed a proof of principle analysis to demonstrate this approach, testing the previously established finding that maternal smoking in pregnancy leads to lower offspring birthweight. Second, we tested for causal effects of maternal smoking on offspring later life outcomes. Finally, we tested for a causal effect of grandmother's smoking on grandchild's birthweight.

Methods

Study population

Our study was conducted using UK Biobank, a population-based cohort of more than 500 000 men and women in the UK. The UK Biobank received ethical approval from the research ethics committee (REC reference for UK Biobank 11/NW/0382) and participants provided written informed consent.

This study collected a large and diverse range of data from physical measures, questionnaires and hospital episode statistics.³² Of 463 013 participants of European descent with genetic data passing initial quality control (i.e. genetic sex same as reported sex, XX or XY in sex chromosome and no outliers in heterozygosity and missing rates),³³ 289 684 participants (54% women) of White British descent were eligible for inclusion in our analyses (Supplementary Figure 1, available as Supplementary data at *IJE* online). We refer to the UK Biobank participants as generation one (G1), and their mothers and offspring as G0 and G2, respectively.

Genetic IV for maternal (G0) smoking

The rs16969968 SNP located in *CHRNA5* has been robustly associated with smoking heaviness.²⁶ Ideally, we would use the maternal rs16969968 as an IV for the heaviness of maternal smoking, but in UK Biobank parental genetic data are not available. Hence, we used rs16969968 of the UK Biobank participants (G1) as a proxy for that of

their mothers, coded as the number of smoking heaviness-increasing alleles.

Smoking phenotypes

We used participants' answers to the question 'Did your mother smoke regularly around the time when you were born?' as a proxy for G0 smoking during pregnancy. Participants were also asked to report their smoking status (current/former/never). We derived a binary ever versus never measure of smoking status by combining current and former smokers. We derived a measure denoting whether G1 women who had at least one live birth smoked during the pregnancy of their first child. G1 women who gave birth to their first child at age b were assigned as non-smokers in pregnancy if they: (i) never smoked; (ii) currently smoked but started smoking when $b + 1$ years old or later; (iii) formerly smoked but started smoking when $b + 1$ years old or later; or (iv) formerly smoked but stopped smoking when $b - 2$ years old or earlier. Smokers included: (i) current smokers who started smoking when $b - 1$ years old or earlier; and (ii) former smokers who started smoking when $b - 1$ years old or earlier but stopped smoking when $b + 1$ years old or later. As these ages were recorded as whole numbers of years, it was not always possible to determine whether a woman was a smoker in pregnancy (illustrated in Supplementary Figure 2, available as Supplementary data at *IJE* online).

Outcomes in participants (G1)

We used baseline data measured at the UK Biobank initial assessment centre. Anthropometric traits included participants' birthweight (kg, self-reported), standing height (cm) and BMI (kg/m^2 , constructed from standing height and weight). To assess lung function, forced vital capacity (L) and forced expiratory volume in 1 s (L) were measured by spirometry. Participants reported whether they had had asthma via the question 'Has a doctor ever told you that you have had any of the following conditions?' (with an option of asthma).³⁴ Systolic and diastolic blood pressure (mmHg) were measured twice using a digital monitor, or a manual sphygmomanometer if the digital monitor could not be employed, and we took the average of the two readings. Female participants reported their age at menarche. We derived years of education based on qualifications achieved by participants, as described previously.³⁵ We included follow-up data of a subset of participants to define intelligence and depression/anxiety. Fluid intelligence score was generated as an unweighted sum of the number of correct answers given to 13 questions, and we used the earliest score if we had data at multiple time points.³⁶ We defined

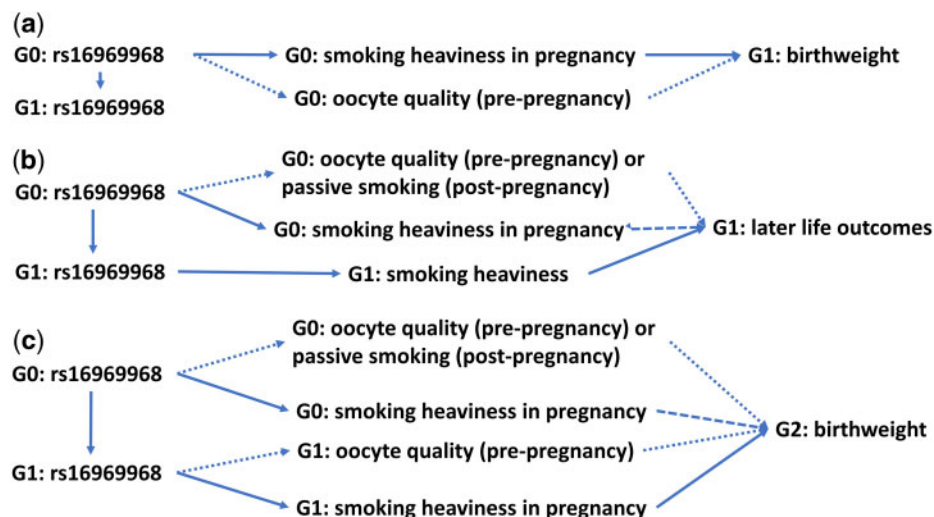


Figure 1. Directed acyclic graphs (DAGs) of this study. Generation (G0): UK Biobank participants' mothers; G1: UK Biobank participants themselves; G2: first offspring of UK Biobank participants. (a) Assessing the effect of G0 smoking heaviness on G1 birthweight: We used G1 rs16969968 as a proxy for G0 rs16969968 and stratified on G0 smoking status in pregnancy. G1 rs16969968 could have no effect on G1 birthweight via G1 smoking heaviness, since G1 cannot smoke before they were born. Maternal smoking outside pregnancy might influence the outcome, e.g. via oocyte quality, causing an alternative path between rs16969968 and G1 birthweight (shown as dotted \rightarrow) (b) Assessing the effect of G0 smoking on G1 later life outcomes: Besides the paths described in (a), G1 rs16969968 could influence the outcomes via G1 smoking heaviness. To estimate the effect of G0 smoking heaviness in pregnancy (shown as dashed \rightarrow), we need to block this by further stratifying on G1 smoking status. (c) Assessing the effect of G0 smoking on G2 birthweight: besides the paths described in (a), G1 rs16969968 could influence G2 birthweight via G1 smoking heaviness in pregnancy. To estimate the effect of G0 smoking heaviness in pregnancy (shown as dashed \rightarrow), we need to block this by further stratifying on G1 smoking status in pregnancy. G1 pre-pregnancy smoking might influence G2 birthweight (shown as dotted \rightarrow). See further DAGs in [Supplementary Figure 3](#) (available as [Supplementary data](#) at *IJE* online) illustrating potential sources of bias due to conditioning on a collider.

depression/anxiety cases as participants that either answered 'Yes' to 'Have you ever seen a general practitioner (GP) for nerves, anxiety, tension or depression?' or to 'Have you ever seen a psychiatrist for nerves, anxiety, tension or depression?', or had hospital episode coded using ICD-10.³⁷ Happiness was assessed via a question – 'In general how happy are you?', with six categories ranging from 'extremely happy' to 'extremely unhappy'.

Outcomes in participants' offspring (G2)

The female participants with at least one live birth were asked to report their first child's birthweight. Male participants were not asked to report the birthweight of their offspring.

Statistical analyses

Proof of principle analysis implies testing the causal effect of maternal (G0) smoking heaviness in pregnancy on participants' (G1) birthweight

In this proof of principle analysis, we seek to replicate the finding, previously established using $G \times E$ MR and many other methods,⁶ that heavier maternal smoking causes lower offspring birthweight. We use our proxy $G \times E$ approach, where participants' (G1) genotype is used

as a proxy for their mothers' (G0) genotype. To assess whether G1 rs16969968 affects G1 birthweight via G0 smoking in pregnancy, we stratified our G1 sample by G0 smoking status during pregnancy, and then tested the associations of rs16969968 with birthweight in each stratum using multivariable linear regression. Since birth precedes smoking initiation, participants' genotype cannot affect birthweight through their own smoking heaviness, which means we do not need to consider smoking status of participants (Figure 1a). We included participants' sex as a covariate to reduce variation in their birthweight, and the first 10 principal components to control for population stratification. We assumed an additive genetic effect and identified the strength of interaction between strata using Cochran's Q test for heterogeneity.

Testing for causal effects of G0 smoking in pregnancy on G1 later life outcomes

We use the proxy $G \times E$ MR approach to test for causal effects of maternal (G0) smoking heaviness on offspring (G1) height, BMI, lung function, asthma, blood pressure, age at menarche, education, intelligence, depression/anxiety and happiness. In contrast to our proof of principle example where participants (G1) smoking in adulthood cannot influence their birthweight, participants'

rs16969968 could affect these outcomes via both maternal (G0) and participants' (G1) smoking heaviness (Figure 1b). To assess whether rs16969968 may affect these outcomes via maternal versus participants' smoking, we stratified on both maternal and participants' smoking status. In each stratum, we examined associations of rs16969968 with height, BMI, lung function, blood pressure, age at menarche, education and intelligence using linear regression, with asthma and depression/anxiety using logistic regression and with happiness using ordinal logistic regression. We included participants' age at baseline, sex and the first 10 genetic principal components as covariates.

Height and age at menarche manifest around the time of puberty, such that participants' own smoking can only affect these if they started smoking before these outcomes are determined. We conducted sensitivity analyses for these outcomes stratifying G1 participants according to whether they were ever smokers before achieving their adulthood height (assuming age at 17 for men and 15 for women³⁸) or their age at menarche.

Testing for causal effects of G0 smoking in pregnancy on grandchild's (G2) birthweight

To test for a causal effect of participants' mothers' (G0) smoking on birthweight of participants' offspring (G2), we stratified G1 women based on their own and their mothers' smoking status during pregnancy, as rs16969968 could affect G2 birthweight through both G0 and G1 smoking heaviness (Figure 1c). Within each stratum, we assessed associations of G1 rs16969968 with G2 birthweight using linear regression, adjusting for the first 10 genetic principal components. We estimated the strength of interaction between G0 smokers and G0 non-smokers within each G1 stratum. We also calculated a difference³⁹ in those associations between G0 smokers and G0 non-smokers within each G1 stratum, and estimated the strength of interaction between two differences to investigate whether G1 smoking status modulates the effect of G1 rs16969968 on G2 birthweight.

Our $G \times E$ MR may be vulnerable to collider bias^{29,40,41} (see details in Discussion), so we tested associations of G1 rs16969968 with G0 and G1 smoking status and potential confounders available in UK Biobank. We performed simulations to compare statistical power of proxy $G \times E$ MR with that of $G \times E$ MR (see Supplementary Methods, available as Supplementary data at *IJE* online). We also tested observational associations of maternal (G0) smoking status with offspring (G1) smoking status and all outcomes for comparison with our MR results. Analyses were performed using R version 3.5.1 (R Foundation for Statistical Computing, Vienna). All code

used to produce the results can be accessed at [<https://github.com/MRCIEU/MR-maternal-smoking>]. Git tag v0.2 corresponds to the version presented here.

Results

Characteristics of participants stratified by sex are shown in Supplementary Table 1 (available as Supplementary data at *IJE* online). Each additional smoking-increasing allele of participants' (G1) rs16969968 was associated with a 1.02 [95% confidence interval (CI): 1.01, 1.03; P -value = 5×10^{-3}] higher odds of their mothers' (G0) smoking in pregnancy, a 0.98 (95% CI: 0.97, 0.99; P -value = 7×10^{-4}) lower odds of being an ever (versus never) smoker themselves, and a 1.06 (95% CI: 1.04, 1.09; P -value = 3×10^{-7}) higher odds that female participants were smokers (versus non-smokers) in their own pregnancy. We found little evidence of an association between rs16969968 and potential confounders, with small associations for participants' age and years of education in some strata (Supplementary Table 2, available as Supplementary data at *IJE* online).

Our proof of principle analysis found that, among participants (G1) whose mothers (G0) smoked in pregnancy, each additional smoking-increasing allele in G1 was associated with a 0.018-kg lower G1 birthweight (95% CI: -0.026, -0.009) after adjustment for covariates (Figure 2). Among participants whose mothers did not smoke in pregnancy, we found little evidence for an association of rs16969968 with birthweight [-0.002 kg (95% CI: -0.008, 0.003)], and we observed heterogeneity between these associations (interaction P -value = 0.004). Estimates were consistent before and after adjustment for covariates.

Figure 3 shows estimates of G1 rs16969968 on the 12 outcomes in the UK Biobank participants (G1). Overall, within each stratum, the estimates were broadly consistent between those whose mothers smoked and those whose mothers did not, except for height among participants who never smoked (all interaction P -values are in Supplementary Table 3, available as Supplementary data at *IJE* online). Each additional smoking-increasing allele was associated with a 0.115-cm lower height (95% CI: -0.200, -0.030) among G1 never smokers whose mothers smoked in pregnancy, but a 0.002-cm lower height (95% CI: -0.057, 0.053) among G1 never smokers whose mothers did not smoke in pregnancy (interaction P -value = 0.029). However, this difference was not observed among G1 ever smokers (Figure 3a). We obtained largely consistent results in sensitivity analyses (Supplementary Figure 3, available as Supplementary data at *IJE* online).

Figure 4 shows results of our analysis testing the effect of grandmother (G0) smoking heaviness on grandchild's

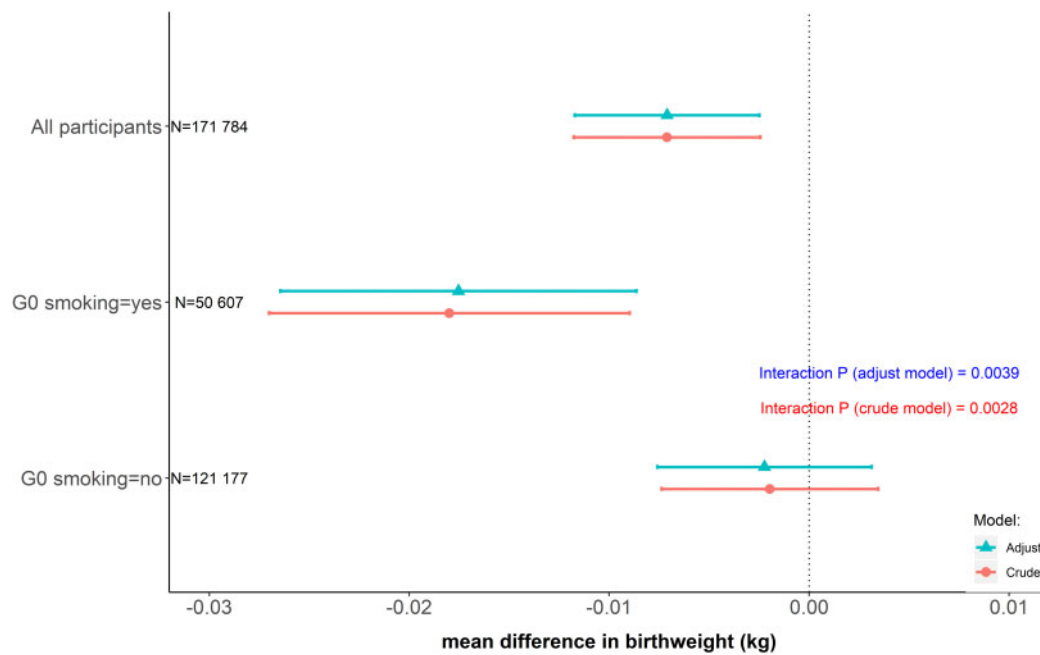


Figure 2. The associations of rs16969968 of UK Biobank participants with their own birthweight by their mothers' smoking status during pregnancy. Generation (G)0: UK Biobank participants' mothers; G1: UK Biobank participants themselves. Estimates are the mean difference of G1 birthweight per each smoking heaviness-increasing allele of rs16969968. In the crude model, we did not adjust for covariates; in the adjusted model, we adjusted for sex of participants and the first 10 genetic principal components. The number of participants was listed for each analysis.

(G2) birthweight. Among mothers (G1) who did not smoke in pregnancy, the association of G1 rs16969968 with grandchild's birthweight in the stratum of grandmother smoking in pregnancy was 0.007 kg/allele (95% CI: -0.005, 0.020) higher than that in the stratum of grandmother not smoking in pregnancy. However, among mothers who smoked in pregnancy, this difference was -0.020 kg/allele (95% CI: -0.044, 0.003). These two differences [pertaining to maternal (G1) smoking strata] were heterogeneous [-0.028 kg per allele (95% CI: -0.055, -0.001); interaction P -value = 0.042].

The directions of observational estimates were consistent with our MR estimates for both participants' (G1) and their children's (G2) birthweight. Our observational analyses also found associations of maternal (G0) smoking in pregnancy with offspring (G1) later life outcomes, where smoking in pregnancy was associated with lower height, higher BMI, poorer lung function, higher risk of asthma, earlier age at menarche, higher blood pressure and poorer cognitive and mental health (Supplementary Table 4, available as Supplementary data at *IJE* online).

Discussion

In this study, we have demonstrated how $G \times E$ MR can be used to test transgenerational causal effects of maternal (G0) smoking heaviness in pregnancy, using offspring (G1) genotype as a proxy for G0 genotype. Our proof of

principle analysis identified an effect of heavier maternal smoking on lower offspring birthweight, consistent with previous studies.²⁻⁶ Our MR study also confirmed previously established causal effects of G1 smoking on their own health, where heavier smoking reduced BMI²⁷ and lead to impaired lung function,⁴² but found little evidence of an effect on asthma risk⁴³ or blood pressure.²⁸

Our tests of effects of maternal (G0) smoking heaviness on offspring (G1) later life health outcomes were not conclusive, given a lack of precision for many of our MR estimates. We found little evidence of an effect on BMI, lung function, asthma, blood pressure, cognition, depression/anxiety or happiness. These findings were consistent with negative control studies for BMI,^{2,8} blood pressure^{19,20} and depression/anxiety,²¹ although our estimation of interactions is not directly quantitatively comparable to their estimation of effects of ever/never smoking or smoking heaviness categories in observational studies. Our MR results found little evidence to support findings from our own and previous observational studies indicating that maternal smoking led to poorer lung function,⁴⁴ higher risk of asthma^{45,46} and lower happiness in offspring.⁴⁷ This may be due to residual confounding in observational associations, or because of low statistical power in MR. Previous studies did not use the same cognition measurement approaches as used in UK Biobank, making our results for this outcome less comparable. We observed lower offspring adulthood height according to maternal smoking in G1 never smokers but not in G1

ever smokers, which could be a chance finding given that we tested multiple outcomes.

We found little evidence of an effect of maternal (G0) smoking in pregnancy on offspring (G1) age at menarche. However, we did find an effect of G1 rs16969968 on age at menarche across strata of both G0 and G1 smoking status, suggesting that rs16969968 may have horizontal pleiotropic effects on age at menarche (e.g. via smoking outside pregnancy). Future MR studies could examine this.²⁵

Our observational results were consistent with previous observational studies^{15–17} by showing a positive association of grandmother’s (G0) smoking in pregnancy with grandchild’s (G2) birthweight after adjusting for mother’s (G1) smoking in pregnancy. Although our $G \times E$ MR was vulnerable to insufficient statistical power, we did find evidence

that female G1 smoking in pregnancy modulates the effect of G0 smoking heaviness in pregnancy on G2 birthweight, consistent with previous observational findings.^{15–17} These results highlight the importance of both grandmother’s and maternal smoking in pregnancy for fetal growth, which could have implications for public health interventions aiming to reduce the prevalence of low birthweight. Two biological mechanisms have been previously suggested to explain the increased growth of offspring (G2) whose grandmothers (G0) did, versus did not, smoke, with non-smoking mothers (G1).⁴⁸ The first theory suggests that mothers (G1) exposed as fetuses to intrauterine smoking [due to grandmother (G0) smoking in pregnancy] would biologically anticipate growth restriction in their child (G2) and thus programme them to grow faster as a pre-adaptation to the

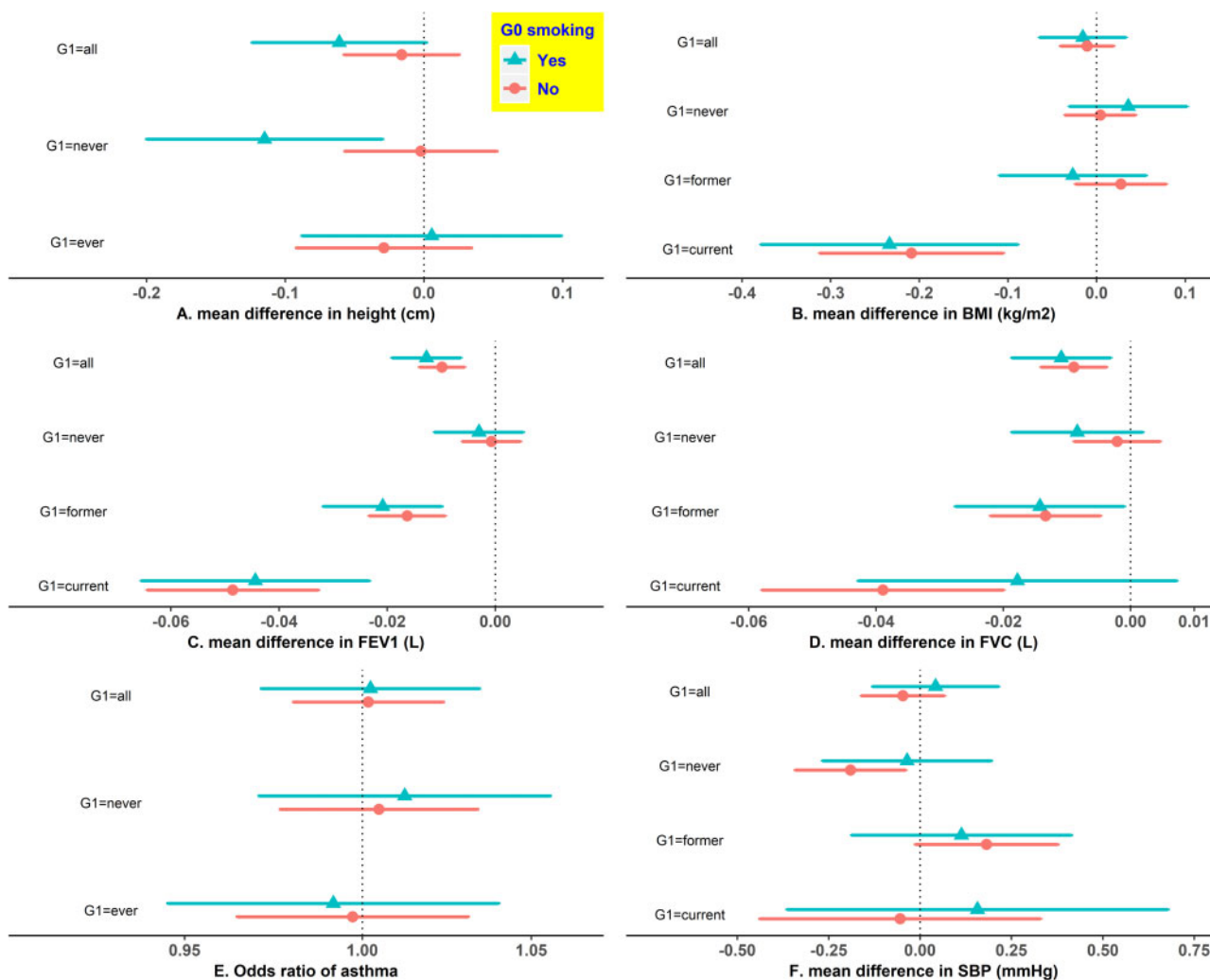


Figure 3. The associations of participants’ rs16969968 with 12 outcomes in UK Biobank participants by their mothers’ smoking status during pregnancy and their own smoking status. Generation (G0): UK Biobank participants’ mothers; G1: UK Biobank participants themselves. Estimates are the mean difference (or change in odds) of G1 outcome per each smoking heaviness-increasing allele of rs16969968. We adjusted for age and sex of participants for outcomes except for menarche, and the first 10 principal components for all 12 outcomes. We combined G1 current and former smokers into ever smokers for height, menarche, education, asthma and happiness, to enlarge sample sizes given smoking cessation may not have a rapid impact on them. BMI, body mass index; DBP, diastolic blood pressure; FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; SBP, systolic blood pressure.

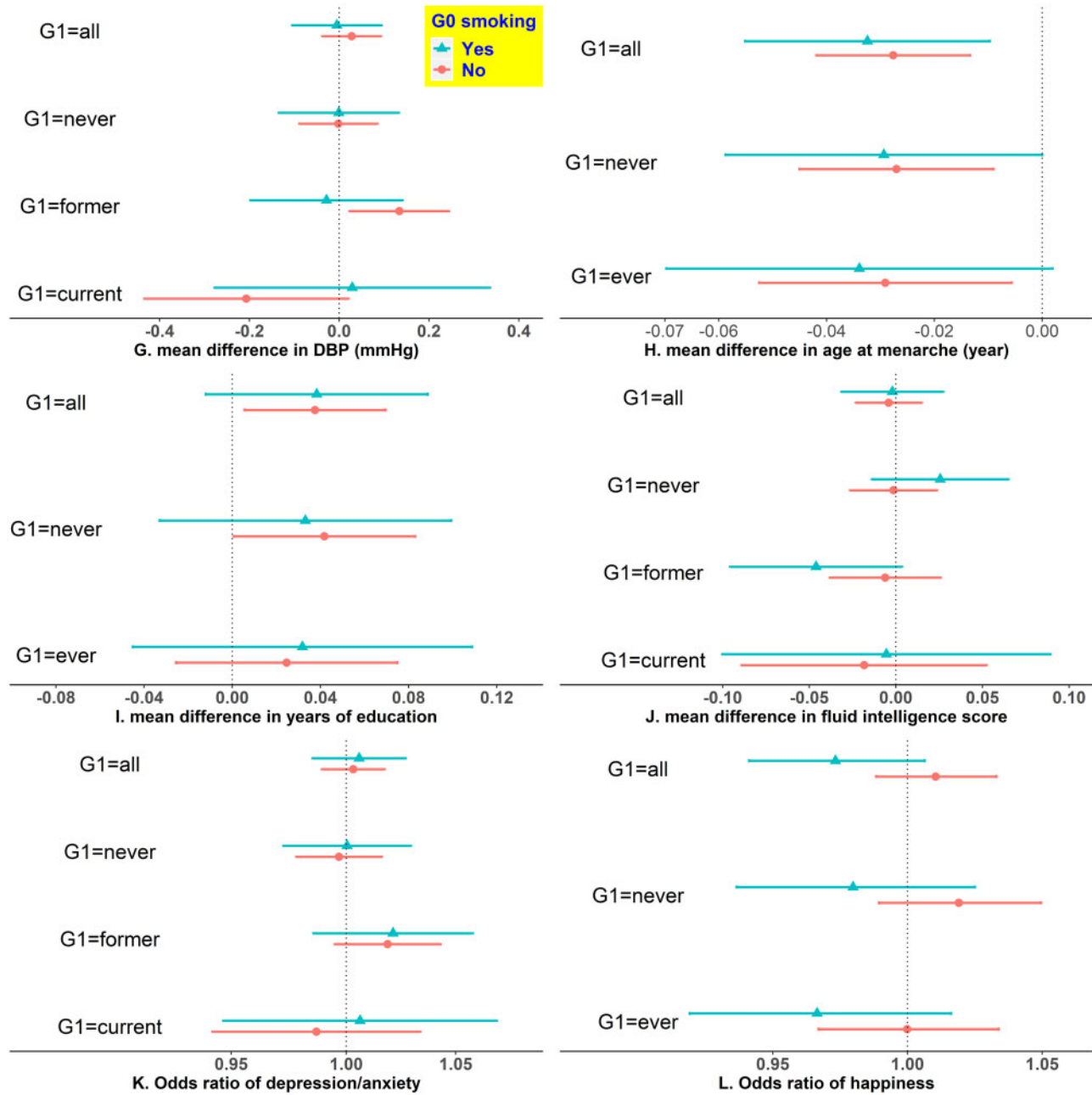


Figure 3. Continued.

anticipated environment.⁴⁸ The second theory suggests that offspring fetal growth may be accelerated by paternal factors, with this normally being constrained by maternal influences.⁴⁹ However, mothers exposed as fetuses to intrauterine smoking could have impaired ability to constrain the paternal impact, creating an imbalance that would cause her child to grow faster.⁴⁸ However, these theories cannot explain the modulation we observed, of grandmother smoking in pregnancy on grandchild’s birthweight, among mothers who did, versus did not, smoke.

We now discuss some limitations of this work. First, our proxy G × E MR used offspring genotype as a proxy

for maternal genotype, and offspring rs16969968 contains 50% information from fathers. This may reduce our statistical power to detect interactions between smoking status strata. We performed simulations to assess the impact of using offspring genotype as a proxy for parental genotype, and found that power was particularly reduced in smaller samples and/or when there were smaller effects of maternal smoking on the outcome (see [Supplementary Figure 5](#), available as [Supplementary data](#) at *IJE* online). Second, our study might be vulnerable to collider bias,^{29,40,41} as we stratified on G0/G1 smoking status (see [Supplementary Figure 3](#), available as [Supplementary data](#) at *IJE* online).

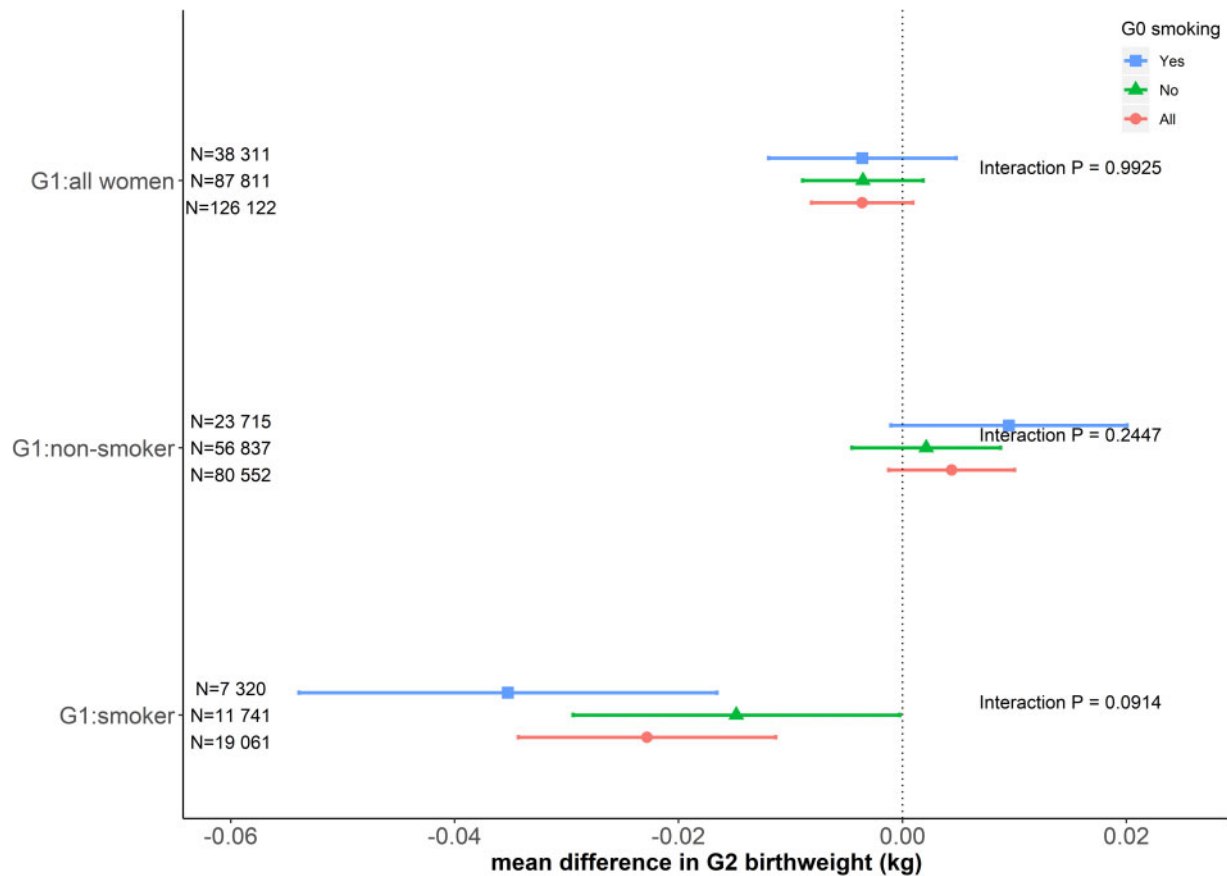


Figure 4. The associations of rs16969968 of UK Biobank women participants with their first child's birthweight by their mothers' and their own smoking status during pregnancy, after adjusting for the first 10 genetic principal components. Generation (G)0: UK Biobank participants' mothers; G1: UK Biobank participants themselves; G2: first offspring of UK Biobank participants. Estimates are the mean difference of G2 birthweight per each smoking heaviness-increasing allele of rs16969968. Interactions are tested between G0 smokers (blue line) and non-smokers (green line) with their *P*-values presented. All women in G1 included G1 smokers, G1 non-smokers and G1 women whose smoking status in pregnancy was missing.

In general, collider bias occurs when a test for the total effect of an instrument (*Z*) on an outcome (*Y*) controls for a third variable (*C*) or a descendant of *C* through stratification/restriction/adjustment, opening up a non-causal pathway between *Z* and *Y*. Additionally, we used an overall sample (UK Biobank) with evidence of selection based on smoking,⁵⁰ and we had some missing data in outcomes which may not have been at random. These may also make our MR estimates vulnerable to collider bias induced by sampling, as described by Hughes *et al.*⁵¹ However, previous simulations^{29,52} and our genetic associations with measured confounders indicated that collider bias may not be large enough to have a meaningful impact on our MR estimates. Third, rs16969968 predicts life course smoking heaviness and not just in pregnancy. Women who smoked in pregnancy will also smoke outside pregnancy. Therefore, the effect of maternal smoking might be via other pathways such as poor oocyte quality (influencing offspring birthweight) or postnatal maternal smoking (e.g. passive smoke exposure)

influencing adulthood outcomes among offspring.⁵³ Using paternal smoking in pregnancy as a negative control² could further validate proxy $G \times E$ MR findings and help to separate intrauterine effects from other effects of parental smoking on offspring outcomes. Unfortunately, such paternal smoking data were not available in UK Biobank.

Fourth, both participants' (G1) and their mothers' (G0) smoking status may be misclassified. Participants were asked to report whether their mother smoked around the time of their birth, and we used this as our measure of G0 smoking in pregnancy. This means that G0 individuals assigned as smokers might have smoked for the duration of their pregnancy or part of their pregnancy, or they might have paused for their pregnancy and restarted smoking shortly after giving birth; whereas G0 individuals assigned as non-smokers might have smoked for a large duration of their pregnancy and stopped sometime before the time of birth. Effects of smoking heaviness in pregnancy may vary according to the duration and pregnancy period during

which a woman smoked. For instance, previous work found that smoking in the first trimester was not associated with lower birthweight in offspring, suggesting that later stages may be more important for fetal growth.^{3,15} Similarly, participants reported their smoking status at baseline, but this may not reflect their smoking status at an important time point for a given outcome. For instance, participants' height and age at menarche can only be affected by their own smoking behaviour if they started smoking before achieving adult height or the onset of puberty. We performed sensitivity analyses for height and age at menarche, using estimates of participants' smoking status before these outcomes. For height, this assumed that men and women achieved their adult height at 17 and 15 years old,³⁸ respectively, as this information was not available in UK Biobank. Fifth, we tested several hypotheses, which increases the probability that our identified associations may be due to chance. Finally, our study may lack statistical power due to small sample sizes in strata and the low power of tests for interactions.⁵⁴ We were unable to account for grandchild's sex in our models assessing the impact of grandmother's smoking in pregnancy since that is unavailable in UK Biobank, which may also reduce our statistical power. MR studies with larger sample sizes and hence greater statistical power are needed to further investigate transgenerational effects of smoking heaviness, together with studies in which both maternal and offspring genotype are known.

We demonstrated how offspring genotype can be used to proxy for maternal genotype, to investigate causal effects of maternal smoking heaviness in pregnancy when maternal genotype is unavailable. We demonstrated our proxy $G \times E$ approach by replicating the previously identified effect of heavier smoking on lower offspring birthweight. We found little evidence of a causal effect of maternal smoking heaviness on offspring's later life outcomes. Finally, we found evidence that the effect of grandmother's smoking in pregnancy on grandchild's birthweight may be modulated by mother's smoking status in pregnancy. Further studies with larger sample sizes are needed to improve statistical power. Maternal drinking heaviness in pregnancy could be another trait to which to apply proxy $G \times E$ MR. More broadly in the field of using MR to investigate transgenerational effects, structural equation modelling could be used to distinguish maternal genetic effects from offspring effects, where either the maternal or offspring genotype is unobserved and modelled as a latent variable.^{55–57}

Supplementary data

Supplementary data are available at *IJE* online.

Acknowledgements

This research has been conducted using the UK Biobank Resource under Application Number 16729 (dataset ID 11148 and 21753).

Funding

This work was supported by the University of Bristol and UK Medical Research Council (grant number MC_UU_12013/1). L.A.C.M. is funded by a University of Bristol Vice-Chancellor's Fellowship. Q.Y. is funded by a China Scholarship Council PhD Scholarship. The funders had no influence on the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Conflict of interest: The authors declare no competing interests.

References

1. Sharp GC, Lawlor DA, Richardson SS. It's the mother!: how assumptions about the causal primacy of maternal effects influence research on the developmental origins of health and disease. *Soc Sci Med* 2018;**213**:20–27.
2. Davey Smith G. Negative control exposures in epidemiologic studies. *Epidemiology* 2012;**23**:350–51; author reply 351–52.
3. Tyrrell J, Huikari V, Christie JT *et al*. Genetic variation in the 15q25 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRN4) interacts with maternal self-reported smoking status during pregnancy to influence birth weight. *Hum Mol Genet* 2012;**21**:5344–58.
4. Rice F, Harold GT, Boivin J *et al*. Disentangling prenatal and inherited influences in humans with an experimental design. *Proc Natl Acad Sci USA* 2009;**106**:2464–67.
5. Davey Smith G. Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? *Basic Clin Pharmacol Toxicol* 2008;**102**:245–56.
6. Krieger N, Davey Smith G. The tale wagged by the DAG: broadening the scope of causal inference and explanation for epidemiology. *Int J Epidemiol* 2016;**45**:1787–808.
7. Howe LD, Matijasevich A, Tilling K *et al*. Maternal smoking during pregnancy and offspring trajectories of height and adiposity: comparing maternal and paternal associations. *Int J Epidemiol* 2012;**41**:722–32.
8. Riedel C, Schonberger K, Yang S *et al*. Parental smoking and childhood obesity: higher effect estimates for maternal smoking in pregnancy compared with paternal smoking—a meta-analysis. *Int J Epidemiol* 2014;**43**:1593–606.
9. Albers L, Sobotzki C, Kuß O *et al*. Maternal smoking during pregnancy and offspring overweight: is there a dose-response relationship? An individual patient data meta-analysis. *Int J Obes* 2018;**42**:1249–64.
10. Brion MJ, Leary SD, Lawlor DA *et al*. Modifiable maternal exposures and offspring blood pressure: a review of epidemiological studies of maternal age, diet, and smoking. *Pediatr Res* 2008;**63**:593–98.
11. Chen Y, Liu Q, Li W *et al*. Association of prenatal and childhood environment smoking exposure with puberty timing: a systematic review and meta-analysis. *Environ Health Prev Med* 2018;**23**:33.

12. Jayes L, Haslam PL, Gratziou CG *et al.* SmokeHaz: systematic reviews and meta-analyses of the effects of smoking on respiratory health. *Chest* 2016;**150**:164–79.
13. Clifford A, Lang L, Chen R. Effects of maternal cigarette smoking during pregnancy on cognitive parameters of children and young adults: a literature review. *Neurotoxicol Teratol* 2012;**34**:560–70.
14. Tiesler CM, Heinrich J. Prenatal nicotine exposure and child behavioural problems. *Eur Child Adolesc Psychiatry* 2014;**23**: 913–29.
15. Ding M, Yuan C, Gaskins AJ *et al.* Smoking during pregnancy in relation to grandchild birth weight and BMI trajectories. *PLoS One* 2017;**12**:e0179368.
16. Miller LL, Pembrey M, Davey Smith G *et al.* Is the growth of the fetus of a non-smoking mother influenced by the smoking of either grandmother while pregnant? *PLoS One* 2014;**9**:e86781.
17. Hypponen E, Davey Smith G, Power C. Effects of grandmothers' smoking in pregnancy on birth weight: intergenerational cohort study. *BMJ* 2003;**327**:898.
18. Macdonald-Wallis C, Tobias JH, Davey Smith G *et al.* Parental smoking during pregnancy and offspring bone mass at age 10 years: findings from a prospective birth cohort. *Osteoporos Int* 2011;**22**:1809–19.
19. Brion MJ, Leary SD, Davey Smith G *et al.* Similar associations of parental prenatal smoking suggest child blood pressure is not influenced by intrauterine effects. *Hypertension* 2007;**49**: 1422–28.
20. Leary SD, Brion MJ, Lawlor DA *et al.* Lack of emergence of associations between selected maternal exposures and offspring blood pressure at age 15 years. *J Epidemiol Community Health* 2013;**67**:320–26.
21. Taylor AE, Carlslake D, de Mola CL *et al.* Maternal Smoking in Pregnancy and Offspring Depression: a cross cohort and negative control study. *Sci Rep* 2017;**7**:12579.
22. Davey Smith G, Lawlor DA, Harbord R *et al.* Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med* 2007;**4**:e352.
23. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;**32**:1–22.
24. Davey Smith G. Use of genetic markers and gene-diet interactions for interrogating population-level causal influences of diet on health. *Genes Nutr* 2011;**6**:27–43.
25. Spiller W, Slichter D, Bowden J *et al.* Detecting and correcting for bias in Mendelian randomization analyses using gene-by-environment interactions. *Int J Epidemiol* 2018;**48**:702–12.
26. Munafo MR, Timofeeva MN, Morris RW *et al.* Association between genetic variants on chromosome 15q25 locus and objective measures of tobacco exposure. *J Natl Cancer Inst* 2012;**104**:740–48.
27. Freathy RM, Kazeem GR, Morris RW *et al.* Genetic variation at CHRNA5-CHRNA3-CHRNB4 interacts with smoking status to influence body mass index. *Int J Epidemiol* 2011;**40**:1617–28.
28. Linneberg A, Jacobsen RK, Skaaby T *et al.* Effect of smoking on blood pressure and resting heart rate: A Mendelian randomization meta-analysis in the CARTA consortium. *Circ Cardiovasc Genet* 2015;**8**:832–41.
29. Millard LAC, Munafo MR, Tilling K *et al.* MR-pheWAS with stratification and interaction: searching for the causal effects of smoking heaviness identified an effect on facial aging. *PLoS Genet* 2019;**15**:e1008353.
30. Kvalvik LG, Skjaerven R, Klungsoyr K *et al.* Can 'early programming' be partly explained by smoking? Results from a prospective, population-based cohort study. *Paediatr Perinat Epidemiol* 2015;**29**:50–59.
31. Liu JZ, Erlich Y, Pickrell JK. Case-control association mapping by proxy using family history of disease. *Nat Genet* 2017;**49**:325–31.
32. Sudlow C, Gallacher J, Allen N *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;**12**:e1001779.
33. Mitchell R, Hemani G, Dudding T *et al.* UK Biobank Genetic Data: MRC-IEU Quality Control, Version 1. 2017. <https://data.bris.ac.uk/data/dataset/3074krb6t2frj29yh2b03x3wxj> (13 May 2019, date last accessed).
34. Zhu Z, Lee PH, Chaffin MD *et al.* A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. *Nat Genet* 2018;**50**:857–64.
35. Okbay A, Beauchamp JP, Fontana MA *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016;**533**:539–42.
36. Hill WD, Marioni RE, Maghzian O *et al.* A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Mol Psychiatry* 2019;**24**:169–81.
37. Howard DM, Adams MJ, Shireli M *et al.* Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat Commun* 2018;**9**:1470.
38. Tanner JM, Whitehouse RH, Takaishi M. Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. I. *Arch Dis Child* 1966;**41**:454–71.
39. Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003;**326**:219.
40. Munafo MR, Tilling K, Taylor AE *et al.* Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol* 2018;**47**:226–35.
41. Taylor AE, Davies NM, Ware JJ *et al.* Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. *Econ Hum Biol* 2014;**13**:99–106.
42. Colak Y, Afzal S, Lange P *et al.* Smoking, systemic inflammation, and airflow limitation: A Mendelian randomization analysis of 98 085 individuals from the general population. *Nicotine Tob Res* 2019;**21**:1036–44.
43. Skaaby T, Taylor AE, Jacobsen RK *et al.* Investigating the causal effect of smoking on hay fever and asthma: a Mendelian randomization meta-analysis in the CARTA consortium. *Sci Rep* 2017;**7**:2224.
44. Magnus MC, Henderson J, Tilling K *et al.* Independent and combined associations of maternal and own smoking with adult lung function and COPD. *Int J Epidemiol* 2018;**47**:1855–64.
45. Silvestri M, Franchi S, Pistorio A *et al.* Smoke exposure, wheezing, and asthma development: a systematic review and meta-

- analysis in unselected birth cohorts. *Pediatr Pulmonol* 2015;50:353–62.
46. Accordini S, Calciano L, Johannessen A *et al*. A three-generation study on the association of tobacco smoking with asthma. *Int J Epidemiol* 2018;47:1106–17.
 47. Menezes AM, Murray J, Laszlo M *et al*. Happiness and depression in adolescence after maternal smoking during pregnancy: birth cohort study. *PLoS One* 2013;8:e80370.
 48. Golding J, Northstone K, Gregory S *et al*. The anthropometry of children and adolescents may be influenced by the prenatal smoking habits of their grandmothers: a longitudinal cohort study. *Am J Hum Biol* 2014;26:731–39.
 49. Frank SA, Crespi BJ. Pathology from evolutionary conflict, with a theory of X chromosome versus autosome conflict over sexually antagonistic traits. *Proc Natl Acad Sci USA* 2011;108(Suppl 2):10886–93.
 50. Swanson JM. The UK Biobank and selection bias. *Lancet* 2012;380:110.
 51. Hughes RA, Davies NM, Davey Smith G *et al*. Selection bias when estimating average treatment effects using one-sample instrumental variable analysis. *Epidemiology* 2019;30:350–57.
 52. Gkatzionis A, Burgess S. Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? *Int J Epidemiol* 2019;48:691–701.
 53. Lawlor D, Richmond R, Warrington N *et al*. Using Mendelian randomization to determine causal effects of maternal pregnancy (intrauterine) exposures on offspring outcomes: sources of bias and methods for assessing them. *Wellcome Open Res* 2017;2:11.
 54. Marshall SW. Power for tests of interaction: effect of raising the Type I error rate. *Epidemiol Perspect Innov* 2007;4:4.
 55. Warrington NM, Freathy RM, Neale MC *et al*. Using structural equation modelling to jointly estimate maternal and fetal effects on birthweight in the UK Biobank. *Int J Epidemiol* 2018;47:1229–41.
 56. Warrington NM, Beaumont RN, Horikoshi M *et al*. Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nat Genet* 2019;51:804–14.
 57. Evans DM, Moen GH, Hwang LD *et al*. Elucidating the role of maternal environmental exposures on offspring health and disease using two-sample Mendelian randomization. *Int J Epidemiol* 2019;48:861–75.

Commentary: Proxy gene-by-environment Mendelian randomization for assessing causal effects of maternal exposures on offspring outcomes

International Journal of Epidemiology, 2020, 1218–1220

doi: 10.1093/ije/dyaa069

Advance Access Publication Date: 1 May 2020



Liang-Dar Hwang ¹ and David M Evans ^{1,2*}

¹The University of Queensland Diamantina Institute, The University of Queensland, Brisbane, Australia and ²Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK

*Corresponding author. University of Queensland Diamantina Institute, 4102, Brisbane, Australia.

E-mail: d.evans1@uq.edu.au

Editorial decision 27 March 2020; Accepted 6 April 2020

There is considerable interest in using the principles of Mendelian randomization (MR) to examine the causal effect of maternal environmental exposures (especially during pregnancy) on offspring health related outcomes. However, using MR in this context is challenging for a number of reasons, including the fact that traditional MR analyses require large numbers of mother–offspring pairs, where at the very least, the mothers and children have been genotyped and phenotype information has been gathered on the children. A major challenge in applying MR in these situations is that across the world there is a paucity of large-scale epidemiological datasets with genome-wide

genotyped mother–offspring pairs. To overcome this problem, Yang *et al.* have developed a new method which they call proxy gene-by-environment MR¹ that builds on previous work in this space.^{2,3}

The fundamental idea behind the authors' method is to use offspring genotype to proxy levels of a maternal environmental exposure of interest in order to investigate a potential causal relationship between levels of the maternal exposure and an offspring outcome. Since mothers and their offspring share half their genes identical by descent, offspring genotype will be correlated with maternal genotype ($r = 0.5$), and can therefore serve as a proxy for maternal