REVIEW ARTICLE

Paediatric and Perinatal Epidemiology **WILEY**

Mendelian randomisation approaches to the study of prenatal exposures: A systematic review

Elizabeth W. Diemer[1] 🄳 | Jeremy A. Labrecque[2] | Alexander Neumann[1,3] | Henning Tiemeier[1,4] | Sonja A. Swanson[2,5]

[1]Department of Child and Adolescent Psychiatry, Erasmus MC, Rotterdam, The Netherlands

[2]Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands

[3]Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, QC, Canada

[4]Department of Social and Behavioral Science, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[5]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Correspondence
Elizabeth W. Diemer, Erasmus MC, Postbus 2040, Rotterdam 3000 CA, The Netherlands.
Email: e.diemer@erasmusmc.nl

Abstract

Background: Mendelian randomisation (MR) designs apply instrumental variable techniques using genetic variants to study causal effects. MR is increasingly used to evaluate the role of maternal exposures during pregnancy on offspring health.

Objectives: We review the application of MR to prenatal exposures and describe reporting of methodologic challenges in this area.

Data sources: We searched PubMed, EMBASE, Medline Ovid, Cochrane Central, Web of Science, and Google Scholar.

Study selection and data extraction: Eligible studies met the following criteria: (a) a maternal pregnancy exposure; (b) an outcome assessed in offspring of the pregnancy; and (c) a genetic variant or score proposed as an instrument or proxy for an exposure.

Synthesis: We quantified the frequency of reporting of MR conditions stated, techniques used to examine assumption plausibility, and reported limitations.

Results: Forty-three eligible studies were identified. When discussing challenges or limitations, the most common issues described were known potential biases in the broader MR literature, including population stratification (n = 29), weak instrument bias (n = 18), and certain types of pleiotropy (n = 30). Of 22 studies presenting point estimates for the effect of exposure, four defined their causal estimand. Twenty-four studies discussed issues unique to prenatal MR, including selection on pregnancy (n = 1) and pleiotropy via postnatal exposure (n = 10) or offspring genotype (n = 20).

Conclusions: Prenatal MR studies frequently discuss issues that affect all MR studies, but rarely discuss problems specific to the prenatal context, including selection on pregnancy and effects of postnatal exposure. Future prenatal MR studies should report and attempt to falsify their assumptions, with particular attention to issues specific to prenatal MR. Further research is needed to evaluate the impacts of biases unique to prenatal MR in practice.

KEYWORDS

instrumental variable, Mendelian randomisation, pregnancy, prenatal

1 | BACKGROUND

Many pregnancy exposures, including maternal nutrition, substance use, and chronic health conditions, are associated with offspring adverse birth outcomes and health across the life course.[1-4] However, mothers who differ in specific prenatal behaviours and traits are also likely to differ in socio-economic status and many other health behaviours, including substance use, exercise habits, diet, social support, and engagement with medical professionals, that could likewise affect or be associated with offspring outcomes.[5] These confounders of the relationship between pregnancy exposures and offspring outcomes are complex constructs that are difficult to measure, as they often relate to an individual's latent tendency to engage in healthy behaviours or to be exposed to risk factors associated with socio-economic position. Therefore, estimates of causal effects of exposures during pregnancy using more traditional analytic techniques that require measuring and adjusting for confounders may be biased.

Instrumental variable analysis proposing genetic variants as instruments, also known as Mendelian randomisation (MR), is an alternative approach to estimate causal effects of exposures on outcomes. In prenatal MR designs, the mothers' genetic variants (eg single nucleotide polymorphisms [SNPs]) are proposed as instruments to examine the effect of an exposure during pregnancy on an offspring outcome. Under specific conditions, MR allows for unbiased estimation of an average causal effect of an exposure on an outcome, even in the presence of unmeasured confounding of the exposure-outcome relationship.[6] An MR study requires an instrument, defined as a variable that meets the following conditions:

1. The instrument Z (ie the genetic variant) must be associated with the exposure X
2. The instrument Z does not affect the outcome Y except through its possible effect on the exposure X (also known as the exclusion restriction)
3. Individuals at different levels of the instrument Z are exchangeable (ie comparable) with regard to counterfactual outcome.

One important implication of condition 3 is that the instrument Z and the outcome Y cannot share any unmeasured causes. A causal structure that meets these requirements is portrayed in Figure 1.

Under these three conditions, investigators can test whether there is an effect of the exposure on the outcome for at least one individual in the study population,[7] and can estimate bounds for the average causal effect.[8,9] In order to obtain a point estimate of an average causal effect, investigators must assume one of a set of additional conditions holds. These conditions vary in strength and plausibility, and some choices of weaker conditions will produce estimates of average causal effects in unidentifiable subgroups of the study population (see Supporting Information for further detail). This choice of condition alters the population to which the estimated effect applies, and a subgroup average causal effect can differ dramatically from the population average causal effect. Therefore,

Synopsis

Study question

How do Mendelian randomisation (MR) studies of prenatal exposures report and attempt to mitigate potential sources of bias?

What's already known

MR is an increasingly popular approach to study effects of the prenatal environment. However, prenatal MR studies are vulnerable to some unique sources of bias.

What this study adds

Prenatal MR studies frequently discuss and attempt to limit biases common in the general MR literature, but rarely discuss problems unique to the prenatal context, including issues related to offspring genotype, the effects of postnatal exposure, and selection on pregnancy.
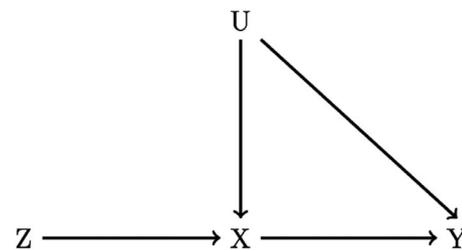
FIGURE 1 Causal Directed Acyclic Graph representing a Mendelian randomisation study where Z is a valid instrument for the effect of X on Y

guidelines for MR analyses recommend explicit reporting of this "fourth" point-identifying condition and the targeted effect estimand.[10-12] Of further note, there are several estimators allowing for relaxation of MR conditions 2 and 3, although these require some alternative assumptions and often the availability of multiple possible instruments.[13-16]

Although the application of MR to pregnancy exposures is growing, to our knowledge, no existing study has examined the frequency of this design, or the assumptions and analytic strategies commonly employed in such applications. As guidelines for MR suggest that the key conditions need to be assessed on a case-by-case basis relative to the study design and research question,[12,17-20] and prenatal MR studies present several unique challenges relative to other types of MR designs,[21,22] it is important to understand how prenatal MR studies report on both study-specific and general challenges to the validity and interpretation of MR results. In addition, by identifying key areas of concern reported by researchers, we

may be able to determine which sources of bias in prenatal MR are in most need of further research. Therefore, the aim of this study was to review the use of MR designs to study the effect of the prenatal environment on offspring outcomes, and to describe the nature and reporting of key potential strengths and weaknesses of the design in this context.

2 | METHODS

To investigate the use of MR in studies of pregnancy exposures, we searched PubMed, EMBASE, Medline Ovid, Cochrane Central, Web of Science, and Google Scholar. Each database was searched from its start date to 14 May 2019. Inclusion in our review required the study met the following criteria: (a) the exposure of interest was a characteristic of the maternal environment that occurred during or proximate to pregnancy, (b) the outcome was assessed in the offspring of the pregnancy, and (c) a genetic variant or genetic variant score was proposed as an instrument and used either as a proxy for an exposure or to conduct an instrumental variable analysis of the effect of exposure on outcome. The inclusion of proxy approaches is especially important for a review of prenatal MR designs, because some early studies did not conceptualise this approach as an application of previously established instrumental variable methods, but rather viewed genetic variants as unconfounded proxies for the exposure of interest. Testing the association between such a genetic variant and an outcome is equivalent to sharp null hypothesis testing in MR and requires the same MR conditions hold.[20] Because birthweight is used both as a characteristic of the offspring and as a proxy for a broad range of characteristics of the prenatal environment, which complicate comparisons to MR analyses of other specific prenatal exposures, we excluded studies using birthweight as an exposure from this review. We also required that the study includes analysis of real data, and we eliminated any duplicate analyses. All studies were independently reviewed by two coauthors (ED & AN), and any disagreements between coauthors were resolved by third author (JL) review and discussion (see Figure 2). Details of the search terms and identified studies are available in the Supporting Information.

Authors extracted data from each included study using a form with open response fields for each data point. Data collected from eligible studies included the study exposure, study outcome, sample size, methodologic approach used, falsification tests and sensitivity analyses performed, and limitations mentioned. For each of the MR conditions, rather than pre-specifying a list of possible types of violations and noting whether a particular article described said violation, reviewers listed all sources of bias described by the article under review that would violate the MR conditions. Although this approach relies on the ability of the reviewer to correctly identify sources of violation that are not explicitly described in the language of instrumental variables (particularly with regard to the fourth assumption), it allows for identification of novel and subject-specific approaches and potential sources of bias, rather than restricting responses to a predefined set of possible violations of the MR conditions. Data were extracted by the first author (ED); to assess accuracy in extraction, five included studies were randomly chosen for independent extraction by a coauthor (JL) (see Supporting Information for details of extraction comparison procedure). Both authors agreed on 56/60 data points (93%) across five articles.
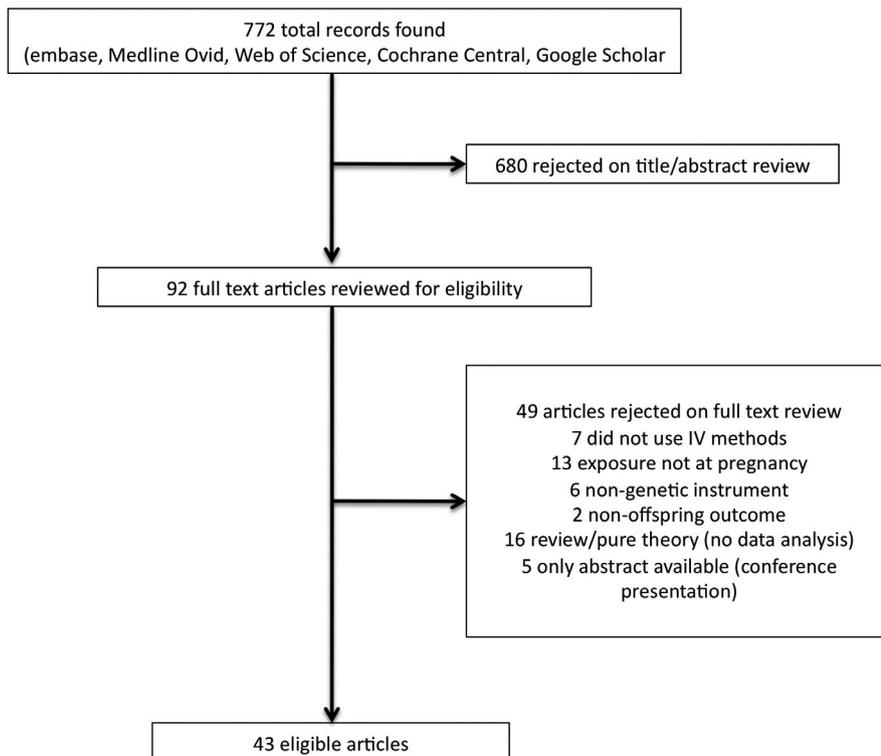

FIGURE 2 Flow chart depicting article eligibility

3 | RESULTS

Initial searches resulted in 772 potentially eligible articles. Of these, 680 articles were excluded based on review of the abstract. Of the 92 articles that underwent full manuscript review, 43 articles met eligibility criteria and were included in this review (Figure 2).

3.1 | Study settings

The included studies covered a wide range of exposures, including alcohol or tobacco use (n = 12, 28%), caffeine use (n = 1, 2%), C-reactive protein (n = 2, 5%), diabetes (n = 4, 9%), thyroid hormone levels (n = 1, 2%), anthropometric traits (n = 8, 19%), placental methylation (n = 1, 2%), haemoglobin levels (n = 3, 7%), blood lipid levels (n = 2, 5%), blood pressure (n = 1, 2%), and micronutrient levels (n = 13, 30%) (Table 1 Column 4). Of the micronutrient studies, six focused on folate, two on vitamin B-12, two on homocysteine, two on vitamin D, and one on polyunsaturated fatty acids. Outcomes of interest included DNA methylation, autoimmune conditions, cognitive development, anthropometric measures (eg adiposity-related outcomes), birthweight, bone density, behavioural disorders, smoking initiation, adverse birth outcomes, orofacial cleft, wheezing, and blood pressure (Table 1 Column 5). The majority (n = 34, 79%) of the studies used data from a birth cohort, with a few studies using case–control designs (n = 4) or cross-sectional data (n = 5). Three studies (7%) used a 2-sample design, in which the association between the proposed instrument and exposure, and between the proposed instrument and outcome, is estimated in independent samples.

The type and number of proposed instruments used varied across included studies. Most (n = 31, 72%) studies proposed only maternal genetic factors as instruments, while the remainder used offspring genetic factors either alone or in tandem with maternal genetic factors. Overall, 19 studies (44%) proposed a single SNP as an instrument, while 24 (56%) used multiple genetic loci.

3.2 | Studies' discussion of key conditions

Eighteen studies (42%) mentioned weak instrument bias, with 10 studies (23%) reporting F-statistics as a measure of proposed instrument strength (range: 0.66-74) (Table S1 Column 11). Seventeen studies (40%) incorporated methods explicitly to limit weak instrument bias into their analysis by leveraging multiple genetic loci as either a genetic risk score or using limited information maximum likelihood and weak instrument robust confidence intervals.[23,24]

Of 15 studies using genetic risk scores, rather than individual SNPs, two explicitly removed SNPs with known pleiotropic effects, that is, SNPs known both to be associated with the exposure and to impact the outcome through paths other than the exposure, from the genetic risk scores. Ten studies (23%) used alternative methods—Egger regression, weighted median regression, and sisVive—which allow for specific types of violations of MR condition 2 under alternative conditions[13-15] (Table S1 Column 16). Ten analyses (23%) controlled for offspring genotype, incorporated offspring genotype into a structural equation model, or used only non-transmitted haplotypes as assumed instruments to mitigate violations of MR condition 2 by offspring genotype.

Twenty-six of the included studies (61%) used some method to avoid violations of MR condition 3 by population stratification, a type of confounding of the proposed instrument-outcome relationship by ancestry group, primarily (n = 19, 44%) via restricting the maternal sample to white European women. Twelve studies (27.9%) included a sensitivity or primary analysis adjusting for GWAS-derived principal components, to limit residual confounding by population stratification. Three studies discussed possible violations of MR condition 3 by assortative mating, a bias resulting from parents selecting mates based on particular characteristics that can result in confounding of the proposed instrument-outcome relationship. One study used linear mixed modelling to mitigate bias resulting from relatedness within the sample.

3.3 | Causal parameters of interest and reporting of additional key conditions

Twenty-one studies (49%) reported proposed instrument-outcome associations only, and 22 (51%) used IV estimation to derive a point estimate of an effect of the exposure on the outcome (Table S1 Column 10). Of the studies that reported such a point estimate, four explicitly reported their estimand of interest (See Supporting Information Sections III-IV for details).

3.4 | Reported sensitivity analyses and falsification tests

While MR conditions 2 and 3 cannot be empirically verified, they can be falsified or indirectly assessed using a variety of techniques.[18,25] However, some of these techniques only detect extreme biases, and, particularly in the case of covariate balance, can be difficult to interpret.[18,25] Three analyses (7%) reported the results of a falsification test (Table 2). One study (2%) estimated a weighting function, and two (5%) used overidentification tests.[26,27] No studies reported instrumental inequalities.[22,28] Twenty-one studies (49%) reported the balance of covariates across levels of their proposed instrument, 17 of which compared this to covariate balance across levels of exposure; no studies used bias or bias component plots to report these comparisons[25] (Table S1, Column 13).

Eleven studies (26%) reported analyses stratified across levels of the exposure or conducted tests of instrument-exposure interaction or interaction between the instrument and a potentially confounded determinant of exposure. One study (2%) stratified across a level of maternal behaviour in which the exposure was expected not to exist,

TABLE 1 Included studies

First author	Year	Proposed instrument(s)	Exposure	Outcome
Allard[62]	2015	2 step[a]: glucose genetic risk score (GRS), methylation GRS	2 step: maternal fasting glucose, methylation	2 step: methylation, cord blood leptin
Alwan[63]	2012	C282Y	Iron	Blood pressure, waist circumference, body mass index (BMI)
Bech[64]	2006	NAT2, CYP1A2, GSTA1	Caffeine	Stillbirth
Bedard[65]	2018	maternal 12 SNP weighted GRS	Haemoglobin	Wheezing, asthma, atopy, low lung function
Bernard[66]	2018	8 FADS variants	Omega 3 and omega 6 polyunsaturated fatty acids	Gestational duration, birthweight, birth length
Binder[67]	2013	MTHFR rs1801133, rs1801131	Folate	Genome-wide methylation
Bonilla[68]	2012	GRS	Fasting glucose, type 2 diabetes	Intelligence quotient (IQ) at age 8
Bonilla[69]	2012	rs492602, rs1801198, rs9606756	Vitamin B12	IQ at age 8
Caramaschi[70]	2017	2 step: rs492602, rs1047781 for vitamin b12; rs5750236, rs1890131 for methylation	2 step: vitamin B12, methylation	2 step: methylation, IQ
Caramaschi[71]	2018	rs1051730	Smoking heaviness	Autism spectrum disorder
Evans[72]	2018	403 SNP GRS	Maternal type 2 diabetes	Birthweight
Geng[73]	2018	35, 25, and 41 SNP GRS	Waist-to-hip ratio adjusted for BMI, hip circumference adjusted for BMI, waist circumference adjusted for BMI	Birthweight, birth length, head circumference
Granell[74]	2008	MTHFR C677T	Folate	Atopy, asthma
Howe[75]	2019	rs1229984	Alcohol	Facial morphology
Humphriss[76]	2013	ADH1B rs1229984	Alcohol	3 composite balance scores (dynamic balance, static balance eyes open, static balance eyes closed)
Hwang[77]	2019	96, 82, and 60 SNP GRS	High-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, triglycerides	Birthweight
Korevaar[78]	2014	GRS	Thyroid-stimulating hormone (TSH), free thyroxine (FT4)	Soluble fms-like tyrosine kinase-1 (sFlt1), placental growth factor (PlGF)
Lawlor[79]	2008	FTO	BMI	Fat mass at age 9-11
Lawlor[21]	2017	GRS	BMI	BMI, fat mass index
Lee[80]	2013	MTHFR C677T	Homocysteine	Birthweight
Lewis[81]	2009	MTHFR C677T	Folate intake	Total weight, total body fat mass, total lean mass
Lewis[82]	2012	10 SNPs in ADH4, ADH1A, AHD1B, ADH7 (rs4699714, rs3763894, rs4148884, rs2866151, rs975833, rs1229966, rs2066701, rs4147536, rs1229984, rs284779)	Alcohol	IQ at age 8
Lewis[83]	2014	GRS based on rs1799945, rs1800562, rs4820268	Iron	IQ at age 8
Mamasoula[84]	2013	MTHFR rs1801133	Folate	Congenital heart disease
Morales[85]	2011	rs1205	C-reactive protein (CRP)	Wheezing, lower respiratory tract infection
Morales[86]	2016	rs1983204, rs344008, rs6795327, rs7637701, rs11929637	Methylation at top-ranked cpg site for placental methylation in smokers	Birthweight

(Continues)

TABLE 1   (Continued)

First author	Year	Proposed instrument(s)	Exposure	Outcome
Murray[87]	2016	GRS including ADH1A rs2866151, rs975833, AHD1B rs4147536, ADH7 rs284779	Alcohol	Conduct problem trajectories (6 measures of Strengths and Difficulties Questionnaire)
Richmond[88]	2016	GRS	BMI	HIF3A methylation
Richmond[89]	2017	GRS	BMI	BMI, fat mass index
von Hinke Kessler Scholder[90]	2014	ADH1B rs1229984	Alcohol	Academic achievement (KS1,KS2,KS3, GCSE)
Shaheen[91]	2014	ADH1B rs1229984	Alcohol	Childhood atopic disease
Steenweg-de Graaff[92]	2012	MTHFR C677T	Folate	Emotional and behavioural score (Child Behavior Checklist)
Steer[93]	2011	MTHFR C677T	Folate	Bone mineral content, bone mineral density, bone area
Taylor[94]	2014	rs1051730	Smoking	Latent class of offspring smoking initiation
Thompson[95]	2019	Separate 7 SNP GRS	Vitamin D, calcium	Birthweight
Tyrrell[96]	2016	GRS	BMI, fasting glucose, diabetes, triglycerides, HDL, blood pressure, vitamin D, adiponectin	Birthweight
Wehby[97]	2011	14 SNPs	Smoking	Birthweight
Wehby[98]	2011	4 SNPS (rs1435252, rs1930139, rs1547272, rs2743467)	Smoking	Orofacial cleft
Wehby[99]	2013	smoking: rs12914385, rs1051730, alcohol: ADH1B rs1229984, BMI: rs8050136	Smoking, alcohol use, obesity	Birthweight
Yajnik[100]	2014	MTHFR rs1801133	Homocysteine	Birthweight
Zerbo[101]	2016	rs3116656, rs2794520	CRP	Autism spectrum disorder
Zhang[102]	2015	GRS	Maternal height	Birth length, birthweight
Zuccolo[103]	2013	rs1229984	Alcohol (1st trimester)	IQ at age 8, educational attainment

[a]2 step Mendelian randomisation designs are a specific subtype of Mendelian randomisation designs proposed to investigate mediation of the relationship between maternal exposures and offspring outcomes by offspring DNA methylation, under additional strong assumptions.[104] In this approach, maternal genetic variants are proposed as instruments for the effect of maternal exposures on offspring methylation across all measured sites. For any methylation sites where a non-null effect was detected for any individual in the population, offspring genetic variants associated with methylation at that site are then proposed as instruments for the effect of methylation at that site on offspring outcomes.

and one study (2%) adjusted for several possible consequences of the proposed instrument and exposure. Because stratifying on or controlling for the exposure or a consequence of the exposure (as in a test of instrument-exposure interaction) can induce collider bias, these analyses will provide a valid falsification test only if there is no confounding of the exposure-outcome relationship, which is extremely unlikely given the typical motivation for conducting an MR analysis.[11]

3.5 | Reported limitations

Thirty-nine studies (91%) discussed versions of potential violations of the MR condition 2, with 30 (70%) describing pleiotropy, 10 (23%) noting possible postnatal effects of the proposed genetic instrument, 14 (33%) discussing possible exposure measurement error, 3 (7%) noting possible preconceptional effects of the proposed genetic instrument on egg quality or maternal characteristics, and 6 (14%) noting their exposure was assumed constant over the course of the pregnancy (Table 3). Thirty-four studies (79%) discussed versions of potential violations of MR condition 3, with most (n = 29, 67% of total) focusing on population stratification. Twenty-eight studies (65%) mentioned low statistical power. Eleven studies (26%) discussed possible selection bias related to missingness of exposure and outcome data. One study (2%) explicitly mentioned selection bias related to the use of a cohort defined by successful pregnancy completion. Sixteen studies (37%) discussed the vulnerability of their analysis to model misspecification resulting from nonlinearity or heterogeneity or violation of proportional hazards. Four studies (9%) noted that they used genetic risk scores weighted based on large GWAS of men and non-pregnant women, which might result in model misspecification when applied to prenatal MR. Of the three studies using two-sample designs, one discussed bias resulting from non-comparability of the samples.

TABLE 2 Falsification approaches and sensitivity analyses reported by included articles

Falsification tests and sensitivity analyses	Per cent studies reporting (n)
Falsification technique	
Overidentification test	5 (2)
Weighting function	2 (1)
Covariate balance	49 (21)
Sensitivity analysis	
Alternative methods (MR-Egger, weighted median, nontransmitted haplotype, SisVive, mode-based estimator)	23 (10)
Pruned GRS	5 (2)
Simulations to evaluate impact of specific type of violation	9 (4)
Adjustment for additional factors	14 (6)
Exposure stratification (would only be valid if no unmeasured confounding of exposure and outcome)	26 (11)

TABLE 3 Possible sources of violation of the MR conditions reported by the included articles

Assumption	Per cent studies reporting (n)
Assumption 1	
Weak instrument bias	42 (18)
Can't prove assumption 1	7 (3)
Winner's curse	2 (1)
Assumption 2	
Pleiotropy	70 (30)
Exposure measurement error	33 (14)
Postnatal effects of genotype	23 (10)
Preconceptional effects of genotype	7 (3)
Exposure assumed constant over pregnancy	14 (6)
Offspring genotype	47 (20)
Assumption 3	
Population stratification	67 (29)
Assortative mating	7 (3)
Residual confounding	16 (7)
Relatedness	2 (1)
Other concerns	
Modelling assumptions	37 (16)
Selection bias—loss to follow-up	26 (11)
Selection on pregnancy	2 (1)
Outcome measurement error	19 (8)
Low power	65 (28)
Limited generalisability	19 (8)
Use of GWAS in nonpregnant adults may be inappropriate	9 (4)
Noncomparable cohort populations (2 sample designs only)	2 (1)

4 | COMMENT

4.1 | Principal findings

The use of MR designs is becoming more frequently applied to study a wide range of types of prenatal exposures and is most often conducted in large, well-characterised birth cohorts. Overall, investigators appear to be aware of possible bias due to pleiotropy and weak associations between proposed instruments and exposures, as well as the low power of MR studies, and demonstrate efforts to address the potential impact of these issues. However, some violations of the MR conditions that are more specific to and perhaps more common in prenatal MR, including violation of MR condition 2 by postnatal or preconceptional exposure status, and selection on pregnancy, are rarely mentioned. The fourth condition used to report point estimates is rarely stated.

4.2 | Strengths of the study

This study is, to our knowledge, the first to investigate the use of the prenatal MR design and the possible violations discussed by applications of this design. The use of prenatal MR is increasing, and a clear evaluation of reported and unreported sources of potential bias is a key consideration for future authors and consumers of prenatal MR studies. By using an open-ended extraction strategy, rather than predefining biases of interest, we were able to identify novel sources of bias specific to this setting. This flexible approach enabled reviewers to identify violations of point-identifying assumptions that were not explicitly described in the language of instrumental variables.

4.3 | Limitations of the data

However, this extraction strategy is, by definition, somewhat subjective. Because this approach relies on the expertise of the reviewer, reproducibility may be impacted. However, when data from five articles were independently extracted by a second coauthor, there was a high degree of agreement between reviewers. As with all systematic reviews, it is also possible that our search algorithm was incomplete, and we did not identify all relevant articles. This limitation is especially relevant to early prenatal MR studies, which did not always use the same language to describe their analysis or conceptualise their analysis as an application of instrumental variables.

Our study focused exclusively on reporting and therefore could not determine whether any potential bias meaningfully impacted the results of a particular study. However, key MR conditions are unverifiable, meaning the absence of all potential biases cannot be proven. Given this, MR studies should, whenever possible, attempt

to falsify their assumptions, and provide sensitivity analyses quantifying the impact of possible biases. If the impact of particular bias is believed to be minor, justification of this assumption based on subject matter knowledge is vital to the interpretation of study findings.

4.4 | Interpretation

Violations of MR condition 2 were some of the most noted problems in this review. Pleiotropy, where genetic loci proposed as an instrument affect both the exposure and another maternal factor associated with the outcome, is a well-recognised problem for all MR studies and was mentioned by nearly three-quarters of the studies (70%) in this review. However, several types of violations of MR condition 2 are relatively unique to the prenatal MR design, some of which remain rarely acknowledged. When maternal genetic factors are proposed as instruments, MR condition 2 could be violated if the offspring's own genotype has an effect on the outcome (Figure 3).[21] This type of bias may be especially common in settings where the maternal exposure and offspring outcome are similar, including studies of the effect of maternal pregnancy BMI on offspring BMI.[21] However, this type of bias could also occur in any setting where offspring exposure level might impact the outcome, or where the mechanism by which a genetic variant proposed as an instrument impacts exposure might also impact the outcome. In MR studies of the effect of prenatal micronutrient exposures on offspring outcomes in later life, MR condition 2 would be violated if offspring micronutrient levels after birth also affect the outcome, because offspring genotype likely impacts their micronutrient levels after birth. Some approaches to limit this bias have been proposed, including controlling for offspring genotype, the use of non-transmitted haplotypes, and a specific linear structural equation model. However, both the nontransmitted haplotype approach and controlling for offspring genotype can induce collider bias via paternal genotype, as both condition on offspring genotype. The structural equation modelling approach proposed by Warrington et al[29] avoids this issue, but requires much stronger assumptions regarding linearity and relationships between covariates than conventional MR.[30]

For maternal proposed genetic instruments, if the outcome of interest occurs after birth, MR condition 2 can be violated if the mother's *postnatal* exposure status also affects the offspring (Figure 4).[19] This is because the mother's genes would logically affect her exposure after birth, and the postnatal effect of the exposure creates an open path between the proposed instrument and the outcome not via prenatal exposure. For example, if the exposure of interest impacts the content of the mother's breastmilk, this would violate MR condition 2. That path is particularly relevant for studies of the effects of obesity, diabetes, substance use, and vitamin B12, all of which have been associated with altered breastmilk content.[31-37] In contrast, previous work has not found any association between maternal iron status and breastmilk content.[36] Altered social exposures and parenting behaviours resulting from maternal postnatal exposure status (eg altered socio-economic status or attachment style
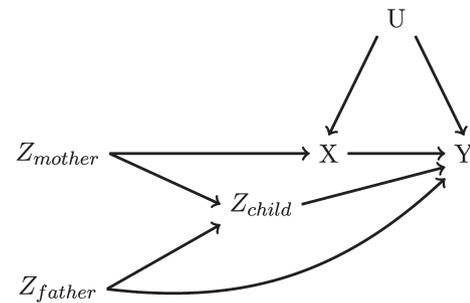

FIGURE 3 Causal Directed Acyclic Graph depicting a maternal genetic loci that violates the MR conditions. Here, offspring genotype (Z_{child}) is an open backdoor path between the proposed instrument (Z_{mother}) and the outcome (Y), violating MR condition 2. However, conditioning on Z_{child} may induce a collider bias if paternal genotype (Z_{father}) is also related to Y, potentially via paternal exposure

resulting from alcohol consumption) may also violate MR condition 2. For studies proposing offspring genetics as instruments, a similar violation can occur if the offspring's genetic factors continue to impact their exposure after birth. For example, as with biases resulting from the causal effect of maternal genotype on offspring genotype, in studies of the effect of micronutrients that propose offspring genetic factors as instruments will be biased if offspring micronutrient levels after birth impact their outcome, as offspring genotype likely continues to affect micronutrient levels after birth. Further, MR condition 2 can be violated if the mother's preconceptional exposure status affects her offspring, through mechanisms like alterations in oocyte quality.

Although an MR estimate of a maternal exposure with postnatal or preconceptional effects could retrieve a valid estimate of the effect of maternal exposure from oocyte formation to outcome measurement, such an approach implies exposures remain the same over several years (in the case of preconceptional effects, from the mother's own gestation to outcome measurement) and do not change as a result of pregnancy, an unreasonable assumption for many exposures of interest.[38] In addition, if the relationship between the proposed genetic instruments and maternal exposure status varies over the course of pregnancy, prenatal MR will produce biased estimates even if the exposure has no postnatal effect.[20,38] Time-varying gene-exposure relationships were not explicitly mentioned in any of the articles reviewed here, though 10 studies mentioned pleiotropy via postnatal or prepregnancy effects as a possible limitation, and six noted the exposure was assumed constant over the course of pregnancy. In settings where postnatal exposure status is believed to substantially affect offspring outcomes, and the gene-exposure relationship varies over time (either before or after birth), prenatal MR with the usual MR estimators will likely be an inappropriate approach, and investigators should consider alternative methods.

Violations of MR condition 3 by population stratification, a problem recognised in the broader MR literature, were well-discussed by studies included in this review (n = 29, 67%).[21,39-42] Violations by selection bias related to participant loss to follow-up, another known problem in MR, were also mentioned by almost a third of studies in this review (n = 11, 26%).[21,39-42] However, because many exposures also negatively impact

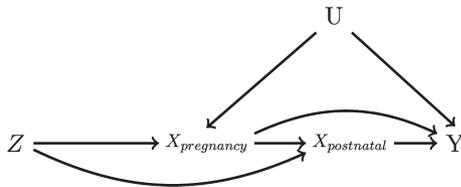**FIGURE 4** Causal Directed Acyclic Graph depicting a maternal genetic locus proposed as an instrument (Z) that violates the MR conditions. Here, Z affects maternal exposure levels both during and after pregnancy, and maternal postnatal exposure also impacts offspring outcomes. Thus, maternal postnatal exposure ($X_{postnatal}$) creates an open backdoor path between Z and the outcome (Y), violating MR condition 2
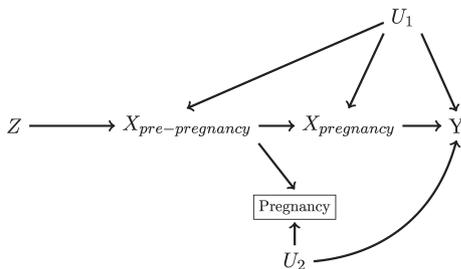

FIGURE 5 Causal Directed Acyclic Graph depicting a maternal genetic locus proposed as an instrument (Z) that violates the MR conditions. Here, the maternal exposure $X_{pre-pregnancy}$ impacts a woman's ability to become pregnant. As outcomes (Y) can only be measured in children of women who successfully conceive and carry a pregnancy to term, a prenatal MR study must necessarily select on pregnancy status, which will generate collider bias in this scenario, violating the MR conditions

fertility or ability to carry a pregnancy to term, prenatal MR studies are also uniquely vulnerable to bias resulting from selecting on successful pregnancy completion (Figure 5), which would result in a violation of the MR condition 3, a limitation mentioned by only one study in this review.[43] This bias could also occur if women with particular substance use and dietary behaviours were less interested in becoming pregnant, or have other lifestyle factors that make it difficult to become pregnant. Previous research suggests that women who are obese are less likely to become pregnant than women who are not obese.[44] Folate status, diabetes, alcohol use, and smoking have been associated with worsened fertility, miscarriage, or stillbirth in experimental animal models and previous observational research.[45-56] Research on becoming pregnant in relation to other exposures included in this review, such as iron status, caffeine use, and C-reactive protein, is less conclusive.[57-60] Because the vast majority of prenatal MR studies are conducted in cohorts recruited based on the presence of a pregnancy, direct correction of estimates using inverse probability weights, a correction approach used in other applications of instrumental variable methods,[43] will rarely be possible. Under specific conditions, the recently proposed MR GENIUS estimator may retrieve unbiased estimates of the causal effect in the presence of selection bias, though this motivation for applying the estimator has not been thoroughly evaluated.[16] As an alternative, authors using prenatal MR might consider using sensitivity analyses informed by previous

research on their exposure and fertility in similar populations to evaluate the robustness of findings to selection bias.[61] However, to this point, no research has examined the magnitude of bias resulting from selection on pregnancy completion in prenatal MR, or optimal bias mitigation and sensitivity analysis strategies in the context of cohorts recruited based on the presence of a pregnancy. It is therefore unclear to what extent prenatal MR studies are biased by selection on pregnancy, and what measures future studies should take to limit or identify this bias.

Some sources of bias in prenatal MR may be particularly difficult to identify via the types of sensitivity analyses and falsification tests used by articles in this review. Comparisons of covariate balance across levels of the instrument and exposure, used by nearly half of the studies in this review, can be difficult to interpret, as even small differences in balance can result in substantial bias.[25] Other methods used in this review, such as overidentification tests, which evaluate the null hypothesis that effect estimates from multiple different instruments are identical, and certain alternative methods allowing for relaxation of MR condition 2, assume that different estimates are not biased in the same way. While this assumption might be reasonable for some forms of horizontal pleiotropy, it will be violated if MR condition 2 or 3 is violated as a result of a shared mechanism like postnatal effects of the exposure, or by selection on pregnancy.[61] Two studies in this review attempted to limit pleiotropy by manually removing SNPs proposed as instruments that had known pleiotropic effects from genetic risk scores. This approach is a useful way of leveraging existing research to identify invalid IVs. However, identifying potentially pleiotropic SNPs in this way requires large GWAS of traits on potential pleiotropic pathways, which may be unavailable for many exposures used in prenatal MR.

While over half of the studies presented point estimates for a causal effect of exposure, few analyses explicitly discuss their estimand (n = 4, 9% of total) or any form of model misspecification (n = 15, 35% of total). As previously stated, because certain choices of weaker modelling assumptions will identify point estimates in different subsets of the population, explicit reporting of investigator assumptions is important to critical evaluation of MR analyses. This is especially true in prenatal MR, where certain subpopulations are not characterised in the same way as conventional MR, and, in the case of certain exposures, including maternal alcohol consumption and smoking, there is evidence that some modelling assumptions are unreasonable (see Supporting Information).

5 | CONCLUSIONS

The use of prenatal MR is especially popular in the study of the effects of adiposity, micronutrient sufficiency, and substance use during pregnancy on offspring health. Because offspring are only directly exposed to maternal genetic factors and certain exposures during gestation, prenatal MR is an appealing method to examine the impact of maternal behaviours on offspring outcomes in the presence of unmeasured exposure-outcome confounding. Authors explicitly discuss and attempt to combat issues that could affect all

MR studies, including population stratification, weak instruments, and certain types of pleiotropy, but much less frequently discuss some of the more specific challenges of prenatal MR designs, such as postnatal effects of the exposure and selection bias related to becoming pregnant. The evaluation of prenatal MR point estimates is also complicated by infrequent reporting of the authors' modelling assumptions and effect of interest, although this pattern has been seen in MR studies and even in other types of instrumental variable analyses more generally.[12,20]

Future studies in this area should include explicit reporting and justification of the authors' assumptions, including those specific to the prenatal context, as well as falsification tests and sensitivity analyses to evaluate the impact of violations of those assumptions. Further research is needed to evaluate how selection bias related to fertility affects prenatal MR estimates, and to determine the best choice of analysis in the presence of violations of the MR conditions in studies of prenatal exposures. Altogether, the relatively frequent reporting of non-specific challenges while underreporting challenges specific to prenatal MR designs may also serve as an important lesson to the developers, teachers, and methodologic collaborators of MR analyses: while published MR applications may be increasingly better at reporting "standard" strengths and limitations of MR studies, critical assessment of the unique challenges of an MR study nonetheless needs to be done on a case-by-case basis.

ACKNOWLEDGEMENT

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

ORCID

Elizabeth W. Diemer https://orcid.org/0000-0002-1701-1414

REFERENCES

1. Fleming TP, Velazquez MA, Eckert JJ. Embryos, DOHaD and David Barker. *J Dev Orig Health Dis*. 2015;6:377-383.
2. Sacks KN, Friger M, Shoham-Vardi I, et al. Prenatal exposure to gestational diabetes mellitus as an independent risk factor for long-term neuropsychiatric morbidity of the offspring. *Am J Obstet Gynecol*. 2016;215(3):380.e1-380.e7.
3. Marques AH, O'Connor TG, Roth C, Susser E, Bjørke-Monsen A-L. The influence of maternal prenatal and early childhood nutrition and maternal prenatal stress on offspring immune system development and neurodevelopmental disorders. *Front Neurosci*. 2013;7:120.
4. Linnet KM, Dalsgaard S, Obel C, et al. Maternal lifestyle factors in pregnancy risk of attention deficit hyperactivity disorder and associated behaviors: review of the current evidence. *Am J Psychiatry*. 2003;160:1028-1040.
5. Smith GD. Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? *Basic Clin Pharmacol Toxicol*. 2008;102:245-256.
6. Hernán MA, Robins JM. *Causal Inference*. Boca Raton, FL: Chapman & Hall/CRC; 2018.
7. Swanson SA, Labrecque JA, Hernán MA. Causal null hypotheses of sustained treatment strategies: what can be tested with an instrumental variable? *Eur J Epidemiol*. 2018;33(8):723-728.
8. Robins J. The analysis of randomized and non-randomized AIDS treatment trails using a new approach to causal inference in longitudinal studies. Health service research methodology: a focus on AIDS. In Sechrest L, Freeman H, Mulley A, eds. *NCHSR, Health Services Research Methodology: A Focus on AIDS*. Washington, D.C.: U.S. Public Health Service; 1989:113-159.
9. Balke A, Pearl J. Bounds on treatment effects from studies with imperfect compliance. *J Am Stat Assoc*. 1997;92:1171-1176.
10. Didelez V, Meng S, Sheehan NA. Assumptions of IV methods for observational epidemiology. *Stat Sci*. 2010; 25(1):22-40.
11. Didelez V, Sheehan NA. Mendelian randomisation and instrumental variables: what can and what can't be done. University of Leicester, Department of Health Science, Technical Report 05. 2005; 2.
12. Swanson SA, Hernán MA. Commentary: How to report instrumental variable analyses (suggestions welcome). *Epidemiology*. 2013;24(3):370-374.
13. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512-525.
14. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40:304-314.
15. Kang H, Zhang A, Cai TT, Small DS. Instrumental variables estimation with some invalid instruments and its application to Mendelian randomization. *J Am Stat Assoc*. 2016;111:132-144.
16. Tchetgen EJT, Sun B, Walter S. The GENIUS approach to robust Mendelian randomization inference. *arXiv*. preprint arXiv:170907779. 2017.
17. Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol*. 2017;14:577.
18. Glymour M, Tchetgen EJT, Robins J. Credible Mendelian randomization studies: approaches for evaluating instrumental variable assumptions. *Am J Epidemiol*. 2012;175:332-339.
19. VanderWeele TJ, Tchetgen EJT, Cornelis M, Kraft P. Methodological challenges in Mendelian randomization. *Epidemiology*. 2014;25:427-435.
20. Swanson SA, Tiemeier H, Ikram MA, Hernán MA. Nature as a trialist? *Epidemiology*. 2017;28:653-659.
21. 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.
22. Diemer EW, Labrecque J, Tiemeier H, Swanson SA. Application of the instrumental inequalities to a Mendelian randomization study with multiple proposed instruments. *Epidemiology*. 2020;31(1):65-74.
23. Finlay K, Magnusson LM. Implementing weak-instrument robust tests for a general class of instrumental-variables models. *Stata J*. 2009;9:398-421.
24. Stock JH, Wright JH, Yogo M. A survey of weak instruments and weak identification in generalized method of moments. *J Bus Econ Stat*. 2002;20:518-529.
25. Jackson JW, Swanson SA. Toward a clearer portrayal of confounding bias in instrumental variable applications. *Epidemiology*. 2015;26:498-504.
26. Hausman JA. Specification tests in econometrics. *Econometrica*. 1978; 46(6):1251-1271.
27. Angrist JD, Imbens GW. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *J Am Stat Assoc*. 1995;90:431-442.

28. Pearl J. On the testability of causal models with latent and instrumental variables. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc.; 1995:435-443.

29. Warrington NM, Freathy RM, Neale MC, Evans DM. Using structural equation modelling to jointly estimate maternal and fetal effects on birthweight in the UK Biobank. *Int J Epidemiol*. 2018;47:1229-1241.

30. VanderWeele TJ. Invited Commentary: Structural equation models and epidemiologic analysis. *Am J Epidemiol*. 2012;176: 608-612.

31. Andreas NJ, Hyde MJ, Gale C, et al. Effect of maternal body mass index on hormones in breast milk: a systematic review. *PLoS ONE*. 2014;9:e115043.

32. Soderborg TK, Borengasser SJ, Barbour LA, Friedman JE. Microbial transmission from mothers with obesity or diabetes to infants: an innovative opportunity to interrupt a vicious cycle. *Diabetologia*. 2016;59:895-906.

33. Young BE, Patinkin Z, Palmer C, et al. Human milk insulin is related to maternal plasma insulin and BMI: but other components of human milk do not differ by BMI. *Eur J Clin Nutr*. 2017;71(9):1094-1100.

34. Giglia R, Binns C. Alcohol and lactation: a systematic review. *Nutr Diet*. 2006;63:103-116.

35. Rowe H, Baker T, Hale TW. Maternal medication, drug use, and breastfeeding. *Pediatric Clinics*. 2013;60:275-294.

36. Allen LH. Multiple micronutrients in pregnancy and lactation: an overview. *Am J Clin Nutr*. 2005;81:1206S-1212S.

37. Koletzko B, Lien E, Agostoni C, et al. The roles of long-chain polyunsaturated fatty acids in pregnancy, lactation and infancy: review of current knowledge and consensus recommendations. *J Perinat Med*. 2008;36:5-14.

38. Labrecque JA, Swanson SA. Interpretation and potential biases of Mendelian randomization estimates with time-varying exposures. *Am J Epidemiol*. 2018;188:231-238.

39. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32:1-22.

40. Palmer TM, Lawlor DA, Harbord RM, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res*. 2012;21:223-242.

41. Frangakis CE, Rubin DB. Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika*. 1999;86:365-379.

42. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res*. 2007;16:309-330.

43. Canan C, Lesko C, Lau B. Instrumental variable analyses and selection bias. *Epidemiology*. 2017;28:396-398.

44. Jokela M, Elovainio M, Kivimäki M. Lower fertility associated with obesity and underweight: the US National Longitudinal Survey of Youth. *Am J Clin Nutr*. 2008;88:886-893.

45. Boots C, Stephenson MD. Does obesity increase the risk of miscarriage in spontaneous conception: a systematic review. *Semin Reprod Med*. 2011;29(06):507-513.

46. Fan D, Liu L, Xia Q, et al. Female alcohol consumption and fecundability: a systematic review and dose-response meta-analysis. *Sci Rep*. 2017;7:13815.

47. Gill J. The effects of moderate alcohol consumption on female hormone levels and reproductive function. *Alcohol Alcohol*. 2000;35:417-423.

48. Bailey BA, Sokol RJ. Prenatal alcohol exposure and miscarriage, stillbirth, preterm delivery, and sudden infant death syndrome. *Alcohol Res Health*. 2011;34:86.

49. Dechanet C, Anahory T, Mathieu Daude JC, et al. Effects of cigarette smoking on reproduction. *Hum Reprod Update*. 2010;17:76-95.

50. Waylen AL, Metwally M, Jones GL, Wilkinson AJ, Ledger WL. Effects of cigarette smoking upon clinical outcomes of assisted reproduction: a meta-analysis. *Hum Reprod Update*. 2008;15:31-44.

51. Feodor Nilsson S, Andersen PK, Strandberg-Larsen K, Nybo Andersen AM. Risk factors for miscarriage from a prevention perspective: a nationwide follow-up study. *BJOG*. 2014;121:1375-1385.

52. Altmäe S, Stavreus-Evers A, Ruiz JR, et al. Variations in folate pathway genes are associated with unexplained female infertility. *Fertil Steril*. 2010;94:130-137.

53. Gaskins AJ, Mumford SL, Chavarro JE, et al. The impact of dietary folate intake on reproductive function in premenopausal women: a prospective cohort study. *PLoS ONE*. 2012;7:e46276.

54. Laanpere M, Altmäe S, Stavreus-Evers A, Nilsson TK, Yngve A, Salumets A. Folate-mediated one-carbon metabolism and its effect on female fertility and pregnancy viability. *Nutr Rev*. 2010;68:99-113.

55. Whitworth KW, Baird DD, Stene LC, Skjaerven R, Longnecker MP. Fecundability among women with type 1 and type 2 diabetes in the Norwegian Mother and Child Cohort Study. *Diabetologia*. 2011;54:516-522.

56. Elbers CC, Onland-Moret NC, Eijkemans MJC, Wijmenga C, Grobbee DE, van der Schouw YT. Low fertility and the risk of type 2 diabetes in women. *Hum Reprod*. 2011;26:3472-3478.

57. Hanson B, Johnstone E, Dorais J, Silver B, Peterson CM, Hotaling J. Female infertility, infertility-associated diagnoses, and comorbidities: a review. *J Assist Reprod Genet*. 2017;34:167-177.

58. Wesselink AK, Wise LA, Rothman KJ, et al. Caffeine and caffeinated beverage consumption and fecundability in a preconception cohort. *Reprod Toxicol*. 2016;62:39-45.

59. Yadegari M, Khazaei M, Anvari M, Eskandari M. Prenatal caffeine exposure impairs pregnancy in rats. *Int J Fertil Steril*. 2016;9:558.

60. Wathes DC, Abayasekara DRE, Aitken RJ. Polyunsaturated fatty acids in male and female reproduction. *Biol Reprod*. 2007;77:190-201.

61. Swanson SA. A practical guide to selection bias in instrumental variable analyses. *Epidemiology*. 2019;30(3):345-349.

62. Allard C, Desgagné V, Patenaude J, et al. Mendelian randomization supports causality between maternal hyperglycemia and epigenetic regulation of leptin gene in newborns. *Epigenetics*. 2015;10:342-351.

63. Alwan NA, Lawlor DA, McArdle HJ, Greenwood DC, Cade JE. Exploring the relationship between maternal iron status and offspring's blood pressure and adiposity: a Mendelian randomization Study. *Clin Epidemiol*. 2012;4:193-200.

64. Bech BH, Autrup H, Nohr EA, Henriksen TB, Olsen J. Stillbirth and slow metabolizers of caffeine: comparison by genotypes. *Int J Epidemiol*. 2006;35:948-953.

65. Bédard A, Lewis SJ, Burgess S, John Henderson A, Shaheen SO. Maternal iron status during pregnancy and respiratory and atopic outcomes in the offspring: a Mendelian randomisation study. *BMJ Open Respir Res*. 2018;5(1):e000275.

66. Bernard JY, Pan H, Aris IM, et al. Long-chain polyunsaturated fatty acids, gestation duration, and birth size: a Mendelian randomization study using fatty acid desaturase variants. *Am J Clin Nutr*. 2018;108:92-100.

67. Binder AM, Michels KB. The causal effect of red blood cell folate on genome-wide methylation in cord blood: a Mendelian randomization approach. *BMC Bioinformatics*. 2013;14:353.

68. Bonilla C, Lawlor DA, Ben-Shlomo Y, et al. Maternal and offspring fasting glucose and type 2 diabetes-associated genetic variants and cognitive function at age 8: a Mendelian randomization study in the Avon Longitudinal Study of Parents and Children. *BMC Med Genet*. 2012;13:90.

69. Bonilla C, Lawlor DA, Taylor AE, et al. Vitamin B-12 status during pregnancy and Child's IQ at Age 8: a Mendelian randomization study in the Avon longitudinal study of parents and children. *PLoS ONE*. 2012;7:e51084.

70. Caramaschi D, Sharp GC, Nohr EA, et al. Exploring a causal role of DNA methylation in the relationship between maternal vitamin B12 during pregnancy and child's IQ at age 8, cognitive performance and educational attainment: a two-step Mendelian randomization study. *Hum Mol Genet*. 2017;26:3001-3013.

71. Caramaschi D, Taylor AE, Richmond RC, et al. Maternal smoking during pregnancy and autism: using causal inference methods in a birth cohort study. *Transl Psychiat*. 2018;8(1):262.

72. Evans DM, Moen GH, Hwang LD, Lawlor DA, Warrington NM. Elucidating the role of maternal environmental exposures on offspring health and disease using two-sample Mendelian randomization. *Int J Epidemiol*. 2019;48(3):861-875.

73. Geng TT, Huang T. Maternal central obesity and birth size: a Mendelian randomization analysis. *Lipids Health Dis*. 2018;17(1):181.

74. Granell R, Heron J, Lewis S, Smith GD, Sterne JAC, Henderson J. The association between mother and child MTHFR C677T polymorphisms, dietary folate intake and childhood atopy in a population-based, longitudinal birth cohort. *Clin Exp Allergy*. 2008;38:320-328.

75. Howe LJ, Sharp GC, Hemani G, Zuccolo L, Richmond S, Lewis SJ. Prenatal alcohol exposure and facial morphology in a UK cohort. *Drug Alcohol Depend*. 2019;197:42-47.

76. Humphriss R, Hall A, May M, Zuccolo L, Macleod J. Prenatal alcohol exposure and childhood balance ability: findings from a UK birth cohort study. *BMJ Open*. 2013;3:e002718.

77. Hwang LD, Lawlor DA, Freathy RM, Evans DM, Warrington NM. Using a two-sample Mendelian randomization design to investigate a possible causal effect of maternal lipid concentrations on offspring birth weight. *Int J Epidemiol*. 2019;48(5):1457-1467.

78. Korevaar TIM, Steegers EAP, Schalekamp-Timmermans S, et al. Soluble Flt1 and placental growth factor are novel determinants of newborn thyroid (dys) function: the generation R study. *J Clin Endocrinol Metab*. 2014;99:E1627-E1634.

79. Lawlor DA, Timpson NJ, Harbord RM, et al. Exploring the developmental overnutrition hypothesis using parental-offspring associations and FTO as an instrumental variable. *PLoS Med*. 2008;5:484-493.

80. Lee HA, Park EA, Cho SJ, et al. Mendelian randomization analysis of the effect of maternal homocysteine during pregnancy, as represented by maternal MTHFR C677T genotype, on birth weight. *J Epidemiol*. 2013;23:371-375.

81. Lewis SJ, Leary S, Smith GD, Ness A. Body composition at age 9 years, maternal folate intake during pregnancy and methyltetrahydrofolate reductase (MTHFR) C677T genotype. *Br J Nutr*. 2009;102(04):493.

82. Lewis SJ, Zuccolo L, Smith GD, et al. Fetal alcohol exposure and IQ at age 8: evidence from a population-based birth-cohort study. *PLoS ONE*. 2012;7:e49407.

83. Lewis SJ, Bonilla C, Brion MJ, et al. Maternal iron levels early in pregnancy are not associated with offspring IQ score at age 8, findings from a Mendelian randomization study. *Eur J Clin Nutr*. 2014;68:496-502.

84. Mamasoula C, Prentice RR, Pierscionek T, et al. Association between C677T polymorphism of methylene tetrahydrofolate reductase and congenital heart disease: meta-analysis of 7697 cases and 13 125 controls. *Circ Cardiovasc Genet*. 2013;6:347-353.

85. Morales E, Guerra S, García-Esteban R. Maternal C-reactive protein levels in pregnancy are associated with wheezing and lower

86. Morales E, Vilahur N, Salas LA, et al. Genome-wide DNA methylation study in human placenta identifies novel loci associated with maternal smoking during pregnancy. *Int J Epidemiol*. 2016;45:1644-1655.

87. Murray J, Burgess S, Zuccolo L, Hickman M, Gray R, Lewis SJ. Moderate alcohol drinking in pregnancy increases risk for children's persistent conduct problems: causal effects in a Mendelian randomisation study. *J Child Psychol Psychiatry*. 2016;57:575-584.

88. Richmond RC, Sharp GC, Ward ME, et al. DNA methylation and BMI: investigating identified methylation sites at HIF3A in a causal framework. *Diabetes*. 2016;65:1231-1244.

89. Richmond RC, Timpson NJ, Felix JF, et al. Using Genetic variation to explore the causal effect of maternal pregnancy adiposity on future offspring adiposity: a Mendelian randomisation study. *PLoS Med*. 2017;14:e1002221.

90. Scholder SVK, Wehby GL, Lewis S, Zuccolo L. Alcohol exposure in utero and child academic achievement. *Econ J*. 2014;124:634-667.

91. Shaheen SO, Rutterford C, Zuccolo L, et al. Prenatal alcohol exposure and childhood atopic disease: a Mendelian randomization approach. *J Allergy Clin Immunol*. 2014;133:225-232.e5.

92. Steenweg-de Graaff J, Roza SJ, Steegers EA, et al. Maternal folate status in early pregnancy and child emotional and behavioral problems: the Generation R Study. *Am J Clin Nutr*. 2012;95(6):1413-1421.

93. Steer CD, Tobias JH. Insights into the programming of bone development from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Am J Clin Nutr*. 2011;94(suppl_6):1861S-1864S.

94. Taylor AE, Howe LD, Heron JE, Ware JJ, Hickman M, Munafò MR. Maternal smoking during pregnancy and offspring smoking initiation: assessing the role of intrauterine exposure. *Addiction*. 2014;109:1013-1021.

95. Thompson WD, Tyrrell J, Borges MC, et al. Association of maternal circulating 25(OH)D and calcium with birth weight: a Mendelian randomisation analysis. *PLoS Med*. 2019;16:e1002828.

96. Tyrrell J, Richmond RC, Palmer TM, et al. Genetic evidence for causal relationships between maternal obesity-related traits and birth weight. *JAMA*. 2016;315:1129-1140.

97. Wehby GL, Fletcher JM, Lehrer SF, et al. A genetic instrumental variables analysis of the effects of prenatal smoking on birth weight: evidence from two samples. *Biodem Soc Biol*. 2011;57:3-32.

98. Wehby GL, Jugessur A, Murray JC, Moreno LM, Wilcox A, Lie RT. Genes as instruments for studying risk behavior effects: an application to maternal smoking and orofacial clefts. *Health Serv Outcomes Res Methodol*. 2011;11:54-78.

99. Wehby GL, Scholder SVHK. Genetic instrumental variable studies of effects of prenatal risk factors. *Biodem Soc Biol*. 2013;59:4-36.

100. Yajnik CS, Chandak GR, Joglekar C, et al. Maternal homocysteine in pregnancy and offspring birthweight: epidemiological associations and Mendelian randomization analysis. *Int J Epidemiol*. 2014;43:1487-1497.

101. Zerbo O, Traglia M, Yoshida C, Heuer LS. Maternal mid-pregnancy C-reactive protein and risk of autism spectrum disorders: the early markers for autism study. *Transl Psychiat*. 2016;6(4):e783.

102. Zhang G, Bacelis J, Lengyel C, et al. Assessing the causal relationship of maternal height on birth size and gestational age at birth: a Mendelian randomization analysis. *PLoS Med*. 2015;12:e1001865.

103. Zuccolo L, Lewis SJ, Smith GD, et al. Prenatal alcohol exposure and offspring cognition and school performance. A Mendelian randomization natural experiment. *Int J Epidemiol*. 2013;42:1358-1370.

respiratory tract infections in the offspring. *Am J Obstet Gynecol*. 2011;204(2):164.e1-9.

104. Relton CL, Davey SG. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol.* 2012;41:161-176.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.