Association of maternal circulating 25(OH)D and calcium with birth weight: A Mendelian Randomisation analysis.

Authors:

William D Thompson,[1,2], Jessica Tyrrell,[1] Maria-Carolina Borges,[2,3] Robin N Beaumont,[1] Bridget A Knight,[1] Andrew R Wood,[1] Susan M Ring,[2,3,4] Andrew T Hattersley,[1] Rachel M Freathy,[1,2†] Debbie A Lawlor.[2,3,5*†]

Affiliations:

1 Institute of Biomedical and Clinical Science, College of Medicine and Health, University of Exeter; 2 MRC Integrative Epidemiology Unit at the University of Bristol; 3 Population Health Science, Bristol Medical School, University of Bristol; 4 Avon Longitudinal Study of Parents and Children, University of Bristol, NIHR Bristol Biomedical Research Centre, 5 Bristol NIHR Biomedical Research Centre

†Joint senor authors who contributed equally

*Correspondence to D.A. Lawlor, MB ChB, FFPH, PhD, MRC Integrative Epidemiology Unit at the University of Bristol, Oakfield House, Oakfield Grove, BS8 2BN, UK. E-mail: d.a.lawlor@bristol.ac.uk

Key Words: 25(OH)D, calcium, birth weight, Mendelian, instrumental variable, ALSPAC, EFSOCH, UK Biobank

Abstract

Background

Systematic reviews of randomized controlled trials (RCTs) have suggested that maternal vitamin D (25(OH)D) and calcium supplementation increase birth weight. However, limitations of many trials were highlighted in the reviews. Our aim was to combine genetic and RCT data to estimate causal effects of these two maternal traits on offspring birth weight.

Methods and Findings

We performed two-sample Mendelian randomisation (MR) using genetic instrumental variables associated with 25(OH)D and calcium that had been identified in genome wide association studies (GWAS; sample 1; N = 122,123 for 25(OH)D and N= 61,275 for calcium). Associations between these maternal genetic variants and offspring birth weight were calculated in the UK Biobank (UKB) (sample 2; N=190,406). We used data on mother-child pairs from two UK birth cohorts (combined N=5,223) in sensitivity analyses to check whether results were influenced by fetal genotype, which is correlated with the maternal genotype (r≈0.5). Further sensitivity analyses to test the reliability of the results included MR-Egger, weighted-median estimator, "leave-one-out" and multivariable MR analyses. We triangulated MR results with those from RCTs, in which we used randomisation to supplementation with vitamin D (24 RCTs, combined N=5,276) and calcium (6 RCTs, combined N=543) as an instrumental variable to determine the effects of 25(OH)D and calcium on birth weight.

In the main MR analysis, there was no strong evidence of an effect of maternal 25(OH)D on birth weight (difference in mean birth weight -0.03g (95%CI: -2.48 to 2.42g, p=0.981) per 10% higher maternal 25(OH)D). The effect estimate was consistent across our MR sensitivity analyses. Instrumental variable analyses applied to RCTs suggested a weak positive causal

effect (5.94g (95% CI: 2.15 to 9.73, p=0.002) per 10% higher maternal 25(OH)D), but this result may be exaggerated due to risk of bias in the included RCTs. The main MR analysis for maternal calcium also suggested no strong evidence of an effect on birth weight (-20g (95%CI: -44 to 5g, p=0.116) per 1 SD higher maternal calcium level). Some sensitivity analyses suggested that the genetic instrument for calcium was associated with birth weight via exposures that are independent of calcium levels (horizontal pleiotropy). Application of instrumental variable analyses to RCTs suggested that calcium has a substantial effect on birth weight (178g (95% CI 121 to 236g, p=1.43 x 10^{-9}) per 1 SD higher maternal calcium level) that was not consistent with any of the MR results. However, the RCT instrumental variable estimate may have been exaggerated owing to risk of bias in the included RCTs. Other study limitations include the low response rate of UK Biobank, which may bias MR estimates, and the lack of suitable data to test whether the effects of genetic instruments on maternal calcium levels during pregnancy were the same as those outside of pregnancy.

Conclusions

Our results suggest that maternal circulating 25[OH]D does not influence birth weight in otherwise healthy newborns. However, the effect of maternal circulating calcium on birth weight is unclear and requires further exploration with more research including RCT and/or MR analyses with more valid instruments.

Author Summary

Why Was This Study Done?

- Birth weight that is lower or higher than average has been associated with poor health outcomes across the lifespan, including infant mortality, cardiovascular disease and type 2 diabetes.

- If we can identify modifiable maternal factors in pregnancy that are causally related to birth weight, we may be able to reduce the number of babies that are born with lower or higher than optimal birth weight. This may in turn help to reduce the associated poor health outcomes.
- Previous studies have suggested that higher maternal 25(OH)D and calcium in pregnancy are associated with higher birth weight. However many of those studies used conventional multivariable regression in observational studies and may be subject to residual confounding. Few of them estimated the size of the effect of either maternal 25(OH)D or calcium levels on birth weight.

What Did the Researchers Do and Find?

- We estimated the effects of maternal gestational 25(OH)D and calcium levels on offspring birth weight using Mendelian Randomization, a method that uses genetic data to overcome certain limitations of traditional observational studies, for example residual confounding.
- We analysed genetic data on 190,406 women from the UK Biobank who reported the birth weight of their first child, along with the results from published studies of genetic associations with 25(OH)D and calcium levels in 122,123 and 61,275 individuals, respectively.
- We checked that the results were not biased by offspring genotype using data from two UK Birth cohorts, the ALSPAC (n=4,576 mother-child pairs) and EFSOCH studies (n=647 mother-child pairs), respectively.
- To strengthen our causal understanding, we triangulated the Mendelian randomization results with findings from randomised controlled trials in which we used randomised status (to vitamin D or calcium supplementation) as an instrumental variable to estimate the effect of 25(OH)D or calcium on birth weight.

- We found no evidence of a strong effect of maternal 25(OH)D on birth weight.
- We found inconsistent evidence of effects of maternal calcium on birth weight.

What Do These Findings Mean?

- Our findings do not support using vitamin D supplementation during pregnancy or pre-conceptually to influence offspring birth weight.
- The effect of calcium on birth weight is still unclear, and needs further investigation in well powered genetic studies and/or well-conducted randomised controlled trials.

Introduction

Infants with lower or higher birth weight (BW) than average are at an increased risk of neonatal mortality and morbidity [1]. BW is also inversely associated with some adverse adult health outcomes, including coronary heart disease[2], type 2 diabetes[3], poor cognitive ability[4] and several types of cancer[5], with most of these associations being linear across most of the BW distribution. BW is an indicator of conditions in utero and may be influenced by modifiable factors in the maternal circulation. For example, there is evidence that higher maternal fasting glucose is causally related to greater fetal growth and higher BW[6,7], which increases the risk of complications during delivery. However relatively little is known about the causal influences of other maternal factors. More evidence is required on how modifying the *in utero* environment might influence BW and associated health outcomes.

Maternal gestational circulating 25(OH)D[8] and calcium[9] may be modifiable risk factors that impact fetal growth and hence BW. Several observational studies using conventional multivariable regression analyses suggest positive associations of maternal 25(OH)D and calcium with infant BW[10-13], however these results might be explained by residual confounding. Systematic reviews of randomised controlled trials (RCTs) of gestational supplementation with vitamin D[14] or calcium[15] have suggested that this supplementation increases BW. However, the authors of the latest vitamin D supplementation systematic review concluded that most of the trials were small and of low quality, and the difference in mean BW was small and unlikely to be of clinical or public health importance[16]. For the calcium supplementation systematic review, the authors noted that for most of the trials there was a low risk of bias based on a score that did not include intention to treat as one of the risk of bias criteria. However, there were high levels of heterogeneity in the results between the trials, bringing into question the clinical importance of calcium supplementation on BW[15].

Mendelian randomisation (MR) is a method in which genetic variants associated with a modifiable exposure are used as instrumental variables to estimate the causal effect of the exposure on an outcome[17]. As genetic variants are fixed at conception, and are generally not associated with classical confounders, MR is less susceptible to bias resulting from reverse causation and residual confounding [17]. We have previously performed a MR study on the effects of maternal adiposity related exposures on BW, finding a positive effect of body mass index (BMI) and blood glucose and an inverse causal effect of systolic blood pressure on BW[6]. In that study, there was evidence of a possible positive causal effect of 25(OH)D on BW, however the confidence intervals were wide and included the null value[6]. A possible causal association of maternal circulating calcium with infant BW was not explored in that study, and to the best of our knowledge there have been no MR studies of that association. The aim of this study was to use MR to explore whether there are causal effects of maternal circulating 25(OH)D and calcium on BW, and if so what the magnitude those effects are. With the release of new UK Biobank data[18], we have a substantially increased the sample size in comparison to the earlier MR study of 25(OH)D, as well as having access to more genetic instruments for 25(OH)D[19], both of which will increase statistical power and hence effect estimate precision. We complement our MR analyses by triangulating results with findings from instrumental variable analyses applied to RCTs of supplementation with vitamin D or calcium[20].

Methods

The analyses plan was developed by RMF, DAL, WDT and MCB, prior to any analyses beginning. It was acted on by WDT and MCB. The analysis plan has not been published but was informally recorded in meeting notes. We made one change to the overall study plans after completing analyses; we undertook risk of bias assessment of the RCTs with a focus on factors that might mean instrumental variable assumptions were violated. This was motivated by differences in results comparing MR to instrumental variables in RCTs, particularly in relation to the effects of calcium on BW. We made one change to the plan following reviewer's comments; we undertook two multivariable MR analyses to adjust for potential confounders. These were: (i) multivariable MR of the association of 25(OH)D with BW adjusting (genetically) for maternal height and (ii) partial multivariable MR of the association of calcium with BW adjusting (genetically) for maternal educational level.

The study design and different data sources are summarized in **Fig 1,** with **Table 1, S1 Text, S2 Text, S3 Text, S4 Text, S1 Table, S2 Table, S3 Table, S4 Table** and **S1 Fig, S2 Fig and S3 Fig** providing more data on each study that has contributed to this paper. Further details, including details of participant consent and ethics approvals, are described in the supplementary material methods. Ethical approval for data extraction from all of the cohorts used in this study was granted from the appropriate authorities (for more details, see **S1 Text**). This study is reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline (S1 Checklist).

Figure 1: Summary of methods and data contributing to this study

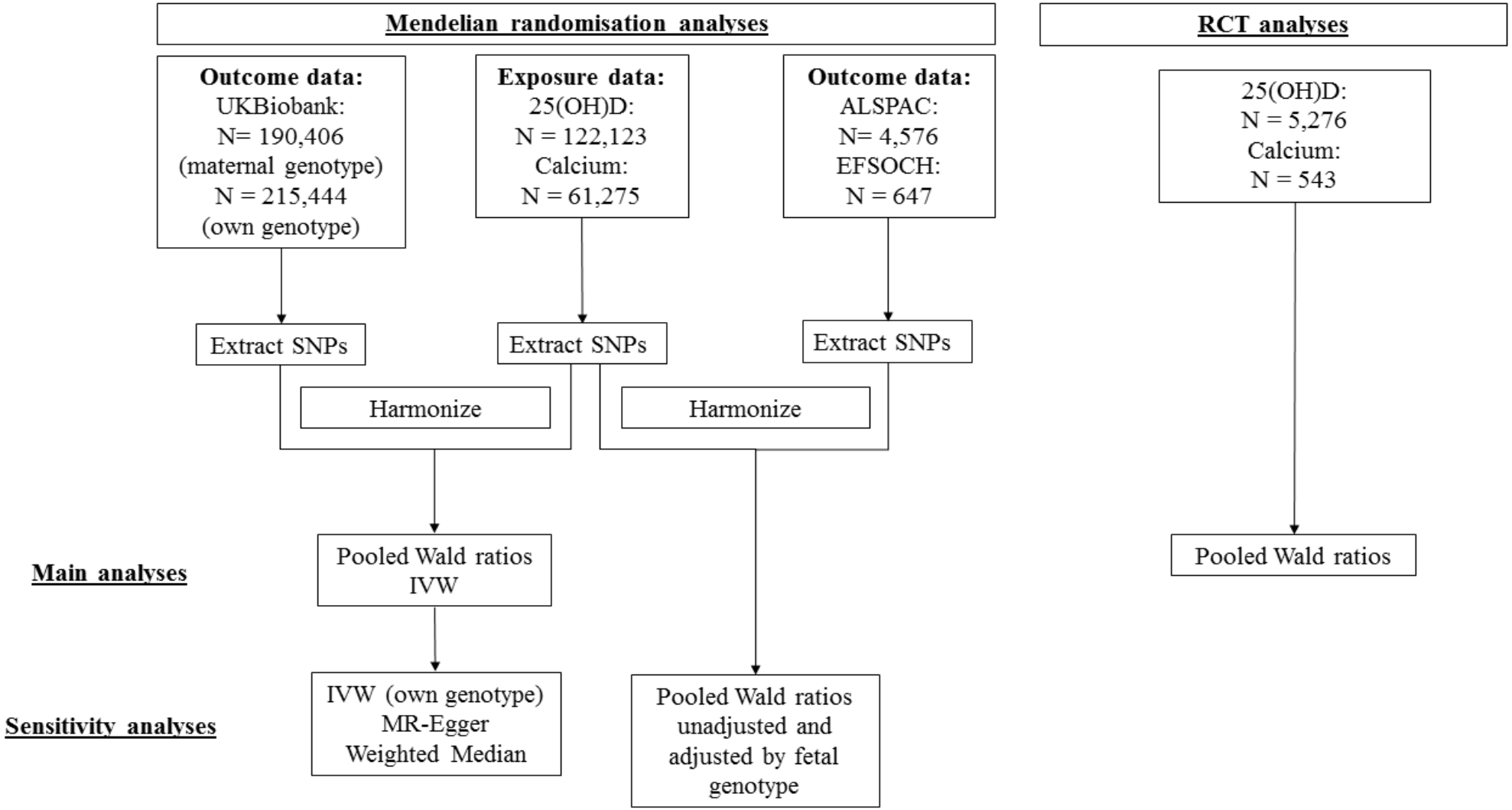


Table 1: Characteristics of the studies used to obtain 25(OH)D/calcium SNPs (genetic instrumental variables) effects on offspring birth weight

	Study	Average/Typical value of Study
Country	UK Biobank	United Kingdom
	ALSPAC	United Kingdom
	EFSOCH	United Kingdom
Offspring years of birth	UK Biobank	1954-2011
	ALSPAC	1991-1993
	EFSOCH	2000-2004
Number of Participants	UK Biobank	190,406
	ALSPAC	4576
	EFSOCH	647
Maternal Age (years)	UK Biobank	25.9 (5.0)
	ALSPAC	29.0 (4.6)
	EFSOCH	30.4 (5.2)
Maternal BMI (kg/m^2)	UK Biobank	27.07 (5.03)
	ALSPAC	22.91 (3.72)
	EFSOCH	24.03 (4.32)
Maternal height (cm)	UK Biobank	162.5 (6.1)
	ALSPAC	164.5 (6.7)
	EFSOCH	165.0 (6.3)
Birth weight (g)	UK Biobank	3227 (476)
	ALSPAC	3495 (471)
	EFSOCH	3514 (475)
Gestational age (weeks)	UK Biobank	NA
	ALSPAC	39.8 (1.3)
	EFSOCH	40.1 (1.2)
Offspring sex (% male)	UK Biobank	NA
	ALSPAC	49
	EFSOCH	52
	UK Biobank	141 (24)

Maternal Systolic Blood Pressure (mmHg)	ALSPAC	133 (13)
	EFSOCH	NA
Mothers smoking (%)[a]	UK Biobank	12
	ALSPAC	15
	EFSOCH	14
Townsend Deprivation Index[b]	UK Biobank	-1.66 (2.86)
	ALSPAC	NA
	EFSOCH	0.23 (3.29)
Educational Attainment: mothers with a university degree (%)	UK Biobank	47
	ALSPAC	14
	EFSOCH	NA
Western Diet[c] (SD)	UK Biobank	-0.087 (0.98)
	ALSPAC	NA
	EFSOCH	NA
Age of assessment for age of first birth (years)	UK Biobank	58.0 (7.8)
	ALSPAC	NA
	EFSOCH	NA

a) In ALSPAC and EFSOCH this is the percentage of women who smoked during pregnancy, for UKB it is the percentage of women who smoke.

b) An area deprivation index that takes summary data on deprivation measures from the census for a defined small geographical area (percentage of households without a motor vehicle, percentage of households with more than one person per room, percentage of households not owner-occupied and percentage of residents who are unemployed), converts them to standard deviation scores across all areas of the UK, then sums them to give a relative area deprivation value, such that a higher score indicates greater deprivation for the area compared to the UK as a

whole[21], in UKB these reflect adult area of residence deprivation based on the postcode provided by the participants at their baseline assessment (aged 40-60 years).

c) Western Diet is a principal component of variation in reported diet in UK Biobank. Variation in diet was measured using a dietary questionnaire[22].

Mendelian Randomisation

We used two-sample summary data MR to explore the effect of maternal circulating vitamin D (25(OH)D) and calcium levels on offspring BW[23]. We used (i) summary data from published genome-wide association studies (GWAS) for the associations of genetic variants (single nucleotide polymorphisms, or SNPs) with 25(OH)D[19,24] or calcium[25] (sample 1) and (ii) summary data for the associations of SNPs with BW from UK Biobank (UKB)[18] (sample 2). Summary data from two UK birth cohorts, the Avon Longitudinal Study of Parents and Children (ALSPAC)[26] and Exeter Family Study of Childhood Health (EFSOCH)[27] were generated for use in sensitivity analyses to explore bias due to fetal genotype. In all studies we excluded participants of non-White European origin (**S2 Text** describes how ethnicity was defined in each study). Following these exclusions we included 190,406 women from UKB who had valid data on BW of first child and GWAS data, 4,576 mother-offspring pairs from ALSPAC and 647 mother-offspring pairs from EFSOCH who all had offspring BW and maternal and offspring GWAS data .

Birth weight and serum 25(OH)D measurement

The UKB is a study of 502,655 participants[28]. Female participants (N=273,495) were also asked to report the BW of their first child. Female participants that reported having a multiple first birth were excluded from our analyses (N=1,364). A total of 216,839 women with a singleton pregnancy for their first child, reported the BW of their first child. Values were reported to the nearest whole pounds, and were converted to kilograms, by multiplying by 0.454, for our analyses. Where women reported the BW of the first child at multiple time points (N=11,353) we used the mean of all measures after excluding any women with a difference of >1kg between any two measures (N=31). We further excluded from the whole sample any women who reported the BW of their first child <2.2kg or >4.6kg (N=6,333).

This was done to reduce bias from reporting errors and, in relation to those <2.2kg to exclude extreme preterm births given we do not have information on gestational age. BW of first child was regressed against the women's reported age at first birth and her UKB assessment centre location to reduce heterogeneity in reporting bias by these characteristics. Residuals from that regression model were then standardized to a mean of 0 and an SD of 1, with the standardized residuals being used in all analyses and final results converted back to grams. The analyses were done on standardized residuals to reduce the amount of computing power needed when doing the initial genome wide association analyses, and so the summary GWAS data that we used for our main analyses was already in this standardized format. UKB participants' (women and men) reports of their own birth weights were used in sensitivity analyses described below and similar methods of exclusion and use of standardized (on age at assessment and centre) residuals, with conversion of results back to grams were used for own BW (see **S3 Text** for further details).

In ALSPAC and EFSOCH, BW was extracted (in grams) from obstetric clinical records at the time of birth, which occurred between 1991 and 1992 (mean (SD) age of mothers 29 (4.6) years) in ALSPAC[26], and between 2000 and 2004 (mean age of mothers 30.4 (5.2) years)in EFSOCH[27].

The GWAS of 25(OH)D and calcium, that we have used in our MR analyses, were undertaken on adult European origin (non-pregnant) women and men. Our MR analyses assume that the magnitude of gene instrument variable-25(OH)D (or calcium) association from those studies are the same in women during pregnancy. We were able to test this for 25(OH)D in ALSPAC, in which 25(OH)D was measured during pregnancy in mothers, using methods that have previously been reported[29] (see **S4 Text**). Neither 25(OH)D nor calcium were measured during pregnancy in UKB or EFSOCH; calcium was not measured in ALSPAC.

Genotyping

For UKB, we analysed data from the May 2017 release of imputed genetic data which has been extensively described elsewhere[30]. Given the reported technical error with non-HRC imputed variants[31], we focused exclusively on the set of ~40M imputed variants from the HRC reference panel.

To account for population structure and relatedness a linear mixed model implemented in BOLT-LMM v2.3[32] was used to perform genome-wide association (GWA) analysis of BW in the UKB sample. Only autosomal single nucleotide polymorphisms (SNPs) which were common (MAF>1%), in Hardy Weinberg equilibrium (p value > 1x10-6), passed QC in all 106 batches and were present on both genotyping arrays were included in the genetic relationship matrix (GRM). For the GWA analyses of BW of the first child (i.e. using the maternal genotype), the genotyping array and genotyping release (interim vs. full) were included as covariates in the regression model. For the GWAS of participants own BW (see below under exploring violation of MR assumptions for the rationale behind these analyses), genotyping array, age at baseline and sex were adjusted for in all models.

In both ALSPAC and EFSOCH, the SNPs used in this study (see below) were taken from genome-wide imputed data that had been completed for both the mothers and their offspring (fetal genotype). In ALSPAC maternal data was obtained from the Illumina 610 Quad Array and fetal data was obtained from the Illumina 550 Quad Array. In EFSOCH maternal and fetal data were obtained from the Illumina Infinium HumanCoreExome-24. For both ALSPAC and EFSOCH, genotype data were imputed against Haplotype Reference Consortium HRC v1.1 reference panel after quality control (MAF >1%, HWE>1×10^{-6}, sex mismatch, kinship errors and 4.56 SD from the cluster mean of any sub-populations cluster)[26,33].

SNP selection, and summary data for SNP-25(OH)D and SNP-calcium associations

We searched for the largest well-conducted GWAS to identify genetic variants (SNPs) that could be used as instrumental variables for circulating 25(OH)D and calcium, and to obtain summary data of genetic instrumental variable (SNP) associations with 25(OH)D and calcium for use in our two-sample MR.

For 25(OH)D, we used summary association results for SNPs identified in two GWAS[19,24], with the largest of these including 79,366 participants from 31 studies in discovery analyses and 42,757 participants from 2 studies in replication analyses. In our main analyses we used seven SNPs, which were not in linkage disequilibrium, from either of the GWAS that had a p-value of 5×10^{-8} in discovery analyses and were replicated. Two of the SNPs discovered in the largest GWAS (rs3755967 and rs17216707) were different from, but in the same loci as, two SNPs identified in an earlier, more commonly used GWAS (rs2282679 and rs6013897, respectively)[34] that have been commonly used as instruments of 25(OH)D in previous MR analyses. We measured linkage disequilibrium between these SNPs in White Europeans (CEU) using LDLink[35,36], and found that both were >0.5, suggesting the SNPs in each pair are tagging the same variant. In additional analyses we separately conducted two analyses with genetic instruments hypothesized to be involved in 25(OH)D synthesis (three SNPs) and 25(OH)D metabolism (two SNPs)[37].

For calcium, we used summary association results for SNPs identified in a GWAS of 39,400 participants [25] from 19 studies in discovery analyses and 21,875 participants from 11 studies as replication. We used seven SNPs, which were not in linkage disequilibrium, that were associated with calcium levels at a p-value of 5×10^{-8} in discovery analysis and were replicated.

Further details of the 25(OH)D and calcium GWAS are provided in **S1 Table**, and a list of the SNPs used in our MR analyses, together with their allele frequencies and per allele associations, for 25(OH)D and calcium are provided in **S2 Table**.

For UKB, the summary results of associations between SNPs and first child BW (maternal genotype) or own BW (own genotype) were extracted from the GWAS results (see above for details on how each GWAS was conducted). For ALSPAC and EFSOCH, individual level SNP data was extracted and summary data were generated using multivariate linear regression of the SNPs against BW (adjusting for gestational age and the child's sex).

To make sure that the outcome data (BW) and exposure data (25(OH)D and calcium) were comparable, the SNPs effects were harmonized to the 25(OH)D/calcium raising alleles using procedures that have previously been described[38].

Statistical Analysis

The main and sensitivity two-sample MR methods are summarized in **Table 2**. In all analyses we estimated the effect of a 10% increase in 25(OH)D on BW in grams, and the effect of 1 SD increase in calcium on BW in grams; these units reflecting the units of 25(OH)D and calcium used in the published GWAS. The value of a 10% increase in 25(OH)D will vary depending on the 'starting point'. In the gestational measures of 25(OH)D in ALSPAC, the median level of 25(OH)D is 61.8 nmol/l, the 25^{th} percentile is 46.1 nmol/l and the 75^{th} percentile is 81.6 nmol/l; this makes a 10% increase from these points equivalent to 6.2 nmol/l, 4.6 nmol/l and 8.2 nmol/l respectively. The accepted range of calcium in a healthy population is between 8.5 mg/dl and 10.5 mg/dl[39], and dividing that range by four we estimated that the SD of calcium is 0.5 mg/dl (~0.3mmol/l).

Table 2: Summary of the four methods used for MR analysis

Name of method	**Wald Ratio (meta-analysis)[40]**	**Inverse-Variance Weighted[41]**	**MR-Egger[41]**	**weighted-median[42]**
Assumption	There is no unbalanced horizontal pleiotropy.	There is no unbalanced horizontal pleiotropy.	That the effect of the genetic instrument is not correlated with any pleiotropic effect of the instrument on the outcome.	Less than 50% of the weight in the analyses come from invalid instruments.
Equation	$Wald\ Ratio = \frac{\beta_{y\|z}}{\beta_{x\|z}}$	$\beta_{IVW} = \frac{\sum_{j=1}^{J} E_j^2 \sigma e_j^{-2} \beta_j}{\sum_{j=1}^{J} E_j^2 \sigma e_j^{-2}}$	$\beta_{MR} = \frac{\rho_j - \beta_0}{E_j}$	$\rho_j = 100(S_j - \frac{w_j}{2})$
Notes on Equation	$\beta_{y\|z}$ is the SNPs effect on the outcome and $\beta_{x\|z}$ is the SNPs effect on the exposure. Wald ratios for each SNP were pooled using fixed-effect meta-analysis with inverse variance weights	βj is the ratio method estimate for each genetic instrument, σej-2 is the standard error of the genetic variants effect on outcome and Ej2 is the genetic instruments effect on the exposure.	β_0 is the intercept, E_j is the genetic instruments effect on the exposure and ρ_j is the genetic instruments effect on the outcome.	Multiple ratio estimates, or β_j, are calculated and the median percentile value is chosen. Importantly, each percentile value is weighted. S_j is the sum of the weights up to the given genetic instrument, ρ_j is the percentile value, and w_j is the weight given to the genetic instrument.

The main effects were calculated in all three studies (UKB, ALSPAC and EFSOCH) using two methods; fixed-effect meta-analysis of Wald ratios[40] of the seven SNPs for 25(OH)D or the seven SNPs for calcium; and the inverse-variance-weighted (IVW) instrumental variable method[41] (**Table 2)**..

Wald ratios were calculated by dividing each SNP's effect on BW by the same SNP's effect on the exposure (25(OH)D or calcium). Standard errors were calculated by dividing the standard error of the SNP's effect on BW by each SNP's effect on the exposure. The ratios for each SNP were then pooled using fixed effect meta-analysis. I^2 and leave-one-out analysis were used to explore between SNP heterogeneity in their MR results (which if present may be due to one or more of the SNPs being an invalid instrumental variable)[43].

For the IVW analysis, linear regression of the weighted (by inverse of their variance) SNPs associations with BW against the SNPs association with maternal circulating 25(OH)D or calcium were performed[41]. In IVW regression analyses, the intercept is forced through zero, making the regression coefficient comparable to the pooled Wald ratio effect estimate. In the presence of heterogeneity, standard errors are larger for IVW compared to pooled Wald ratios. For our analysis we estimated standard errors using a fixed effects model and confidence intervals using a t-distribution.

Exploring possible violations of MR assumptions

Both the Wald ratio method and IVW instrumental variable analysis assume that: (i) that the SNPs being used are robustly associated with maternal circulating 25(OH)D and calcium; ii) the SNPs are not related to confounders of the associations 25(OH)D and/or calcium with BW; and (iii) the SNPs have no effect on BW other than through 25(OH)D and calcium (also known as the exclusion restriction criterion). In MR studies horizontal pleiotropy is a common cause of violation of this assumption. This would occur if our genetic instrument for

25(OH)D or calcium influenced other factors, separately to 25(OH)D or calcium, and these other factors influence birth weight independently of 25(OH)D or calcium. If this were the case, then then estimate of effect that we assumed was due to e.g. 25(OH)D would be the sum of a 25(OH)D effect and the effect on the outcome of any other (pleiotropic) effects. The Wald ratio and IVW approaches complement each other, with determining Wald ratios for each SNP providing an opportunity to explore between SNP heterogeneity and IVW being closely related and comparable to one of our sensitivity analyses (MR-Egger) used to test possible horizontal pleiotropy (**Table 2**).

One possible source of bias is via the fetal genotype[44]. Maternal genetic variants that influence 25(OH)D and calcium will be associated with the distributions of the same genetic variants in the fetus (as mothers may transmit these alleles to their offspring), and if any of these genetic variants affect fetal growth independently of an effect of maternal circulating 25(OH)D/calcium (for example if fetal 25(OH)D or calcium influence fetal growth), there will be an association between maternal SNPs and offspring BW that is not via the mother's gestational 25(OH)D or calcium. We tested this possible source of bias in two ways. First, we adjusted the maternal SNP with offspring BW association for fetal genotype in a total of 5,223 genotyped mother-child pairs from the ALSPAC and EFSOCH studies. The Wald ratio results were estimated separately for each of these two cohorts and then pooled using a fixed-effect meta-analysis. Second, we used IVW to estimate the effect of own 25(OH)D or calcium on own BW (with a total of 215,444 adult women and men reporting their own BW) in UKB, for comparison with the effect estimates of maternal 25(OH)D or calcium on offspring BW. Stronger effects of own (fetal) 25(OH)D/calcium on their BW (compared with maternal gestational circulating levels of these on offspring BW) would suggest the possibility of our main MR analyses of maternal 25(OH)D/calcium levels on offspring BW

being biased by fetal effects (assuming that measurement error in offspring BW and own BW are similar).

We performed three additional tests to investigate possible violations of MR assumptions: MR-Egger[41] and weighted-median estimator[42], which were only used in UKB (which we considered our main analysis cohort and which has adequate statistical power for these analyses), and exploring SNP associations with confounders in UKB, ALSPAC and EFSOCH (further details of these approaches are provided in **S5 Text**).

Instrumental variable analysis applied to RCTs

Instrumental variable methods can be applied to RCTs to quantify the causal effect estimate of the intermediate that the randomised treatment is assumed to influence[45]. For example, here we used RCTs of randomisation to vitamin D supplements to quantify the effect of circulating 25(OH)D on birth weight. This differs from the original aim and analyses of these RCTs which was to determine the causal effect of the supplements. These analyses are similar to Mendelian randomization, except here the instrumental variable is randomized status. This approach has the same underlying assumptions as all instrumental variable analyses, including MR. However, we assume that they key sources of violation of these assumptions will differ between the RCT and MR analyses (e.g. in the RCTs concealment of randomization and intention to treat analyses will be important, whereas in MR horizontal pleiotropy due to linkage disequilibrium will be important). Under this assumption, if results from our MR and RCT analyses are consistent with each other this increases the likelihood that this is the correct causal effect[20].

We used data from recent systematic reviews and meta-analyses of supplements versus placebo, in pregnant women, for both 25(OH)D[16] and calcium[15] to identify individual RCTs that could be used in our instrumental variable analyses applied to RCTs. As different

RCTs used different doses, type of supplement (e.g. vitamin D2 or D3) or mode of delivery (e.g. oral or injection), and because to date relatively little work has used instrumental variables in RCTs to test causal effects, we *a priori* decided we would use a one-sample instrumental variable approach and only include RCTs that provided both the difference in mean BW and difference in mean 25(OH)D/calcium by randomized arm. Each individual RCT in both reviews was searched to identify those that provided difference in mean BW and 25(OH)D/calcium. This resulted in us being able to include 24 (56%) out of 43 RCTs from the most recent pregnancy vitamin D supplementation systematic review (published 2017) and 6 (26 %) out of 23 RCTs included in the most recent calcium supplementation systematic review (2015); all other RCTs had either no information on differences in mean BW or differences in mean 25(OH)D/calcium.

Two of the authors (WDT and DAL) independently extracted the weighted mean differences in 25(OH)D and BW by trial randomised arm (25(OH)D supplement or placebo/other control), together with their respective confidence interval values, from 24 RCTs that presented results for both of these (N = 5,276 mother-offspring pairs)[16] (**S3 Table**). Similarly, for calcium two authors (WDT and M-CB) independently extracted mean differences and confidence intervals for 6 RCTs (543 mother offspring-pairs)[15] (**S4 Table**). Standard errors were calculated by each of the independent abstractors and a third author (RMF), checked consistency between the abstractors with any discrepancy resolved by discussion between four authors, WDT, DAL, M-CB and RMF. One of the calcium RCTs[46] provided a range of the mean difference, which we treated as the 95% confidence intervals when calculating the standard error.

We calculated the Wald ratio estimate for each RCT by dividing the difference in mean BW by the difference in mean 25(OH)D (or calcium) and then pooled the these using a fixed effect meta-analysis and tested for between study heterogeneity using the I^2 statistic and

leave-one-out-analysis. For the vitamin D supplementation RCTs differences in mean 25(OH)D were in nmol/l and for the calcium supplementation RCTs differences in mean calcium were in mg/dl. In order to make the instrumental variable analyses in RCT results comparable with our two-sample MR results, for 25(OH)D we assumed that a 10% difference was equivalent to 6.2 nmol/l (the value for a 10% difference around the median of the distribution of pregnancy 25(OH)D in ALSPAC), and we multiplied the Wald ratio estimates by 6.2 to scale the results so that they represented the difference in mean BW per 10% increase in 25(OH)D. We did the same with calcium, but used 0.5 (0.5 mg/dl being the 1 SD difference value used in our MR analysis) to scale results so that they represented the difference in mean birth weight per 1SD of calcium.

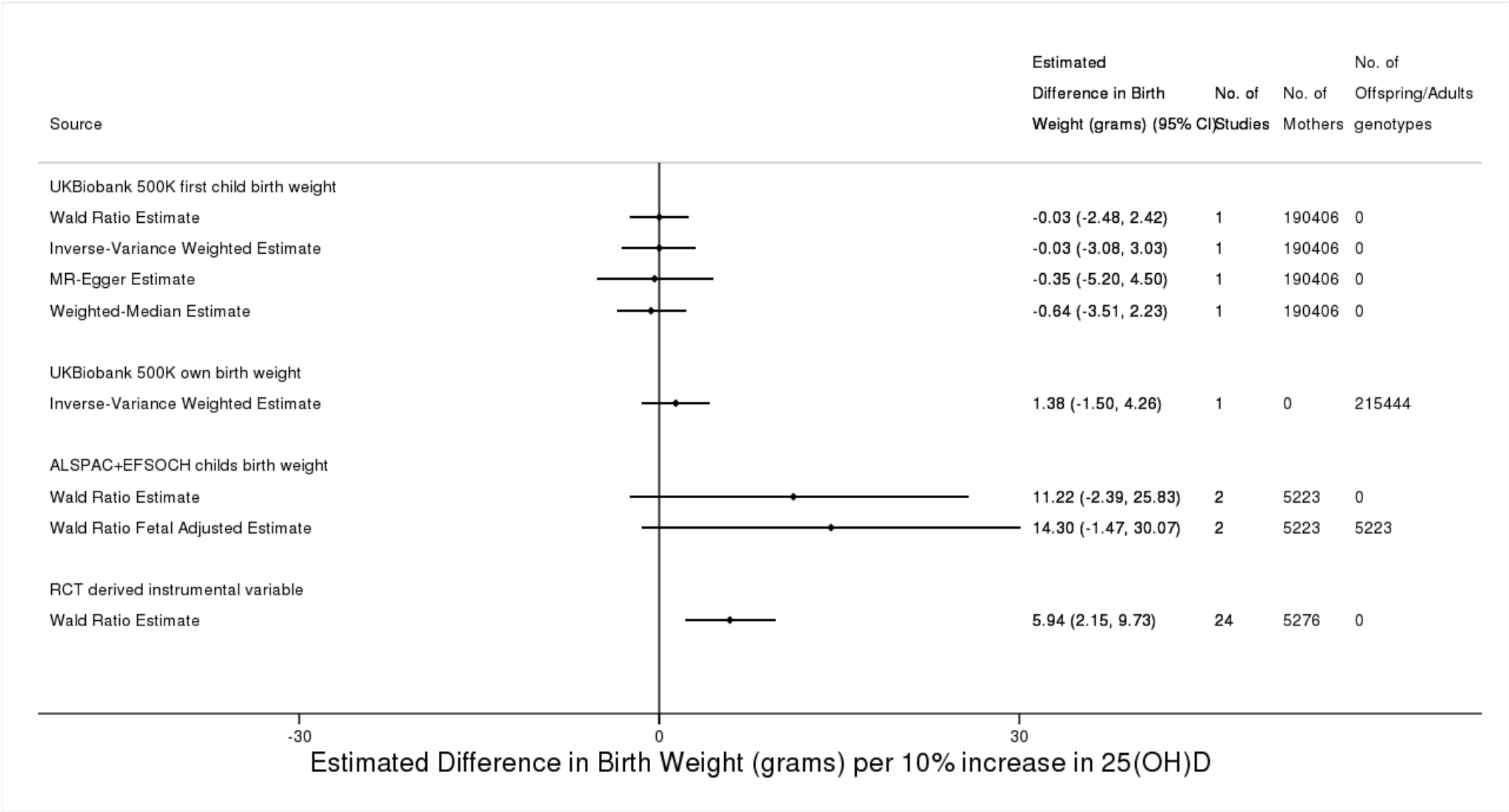
Results

The characteristics of included participants from UKB, ALSPAC and EFSOCH are shown in **Table 1**. The SNP-outcome associations for UKB, ALSPAC and EFSOCH are shown in **S5 Table, S6 Table and S7 Table**.

Mendelian randomisation and instrumental variable analysis in RCTs do not support a clinically important effect of 25(OH)D on birth weight

Our two-sample MR results provide no strong evidence of an effect of maternal circulating 25(OH)D on offspring BW, with consistent null findings in the main Wald ratio and IVW analyses, and also with the MR-Egger and weighted median results using the UKB data (**Fig 2)**. There was no strong evidence for marked heterogeneity between the Wald ratio estimates for each SNP (I^2 = 0.0%) and results were consistent in leave one out analysis with the main results (i.e. no SNPs removed) and with each other (**S4 Fig**). MR effects in ALSPAC and EFSOCH were weakly positive but with wide confidence intervals that included the null value (**Fig 2**). The effect estimates were the same with and without adjustment for fetal genotype in ALSPAC and EFSOCH, and there was no strong evidence that own 25(OH)D influenced own BW in UKB (**Fig 2)**. Our main MR effect estimate in UKB was -0.03g (95% CI, -2.48 to 2.42g, *p*=0.981) per 10% increase in maternal circulating 25(OH)D. However, the RCT instrumental variable effect was 5.94g (95% CI, 2.15 to 9.73, *p*=0.002) per 10% increase in maternal circulating 25(OH)D) (**Fig 2**). There was no strong evidence for marked heterogeneity between the instrumental variable in RCT estimates (I^2=16.2%) and results were consistent, with the main overall result, and with each other, in leave one out analysis (**S5 Fig**).

Figure 2: Causative effect estimates for maternal 25(OH)D on birth weight



The difference between the MR and instrumental variable analyses in RCTs is small, and the confidence intervals for the two results overlap, suggesting that they are statistically consistent. To explore this small difference further we undertook post hoc risk of bias assessment of the 24 RCTs. **S8 Table** summarises the results of this risk of bias assessment. Most studies were small, with the numbers randomised being between 16 and 1134 and only three of the 24 RCTs including more than 200 participants. Where it was reported (16 of the 24 trials) loss to follow-up ranged from 1 to 31%. Only two of the RCTs (12% of participants) had definitely undertaken intention to treat analyses and twelve (54% of studies) had used random sequence and concealed allocation to treatment groups.

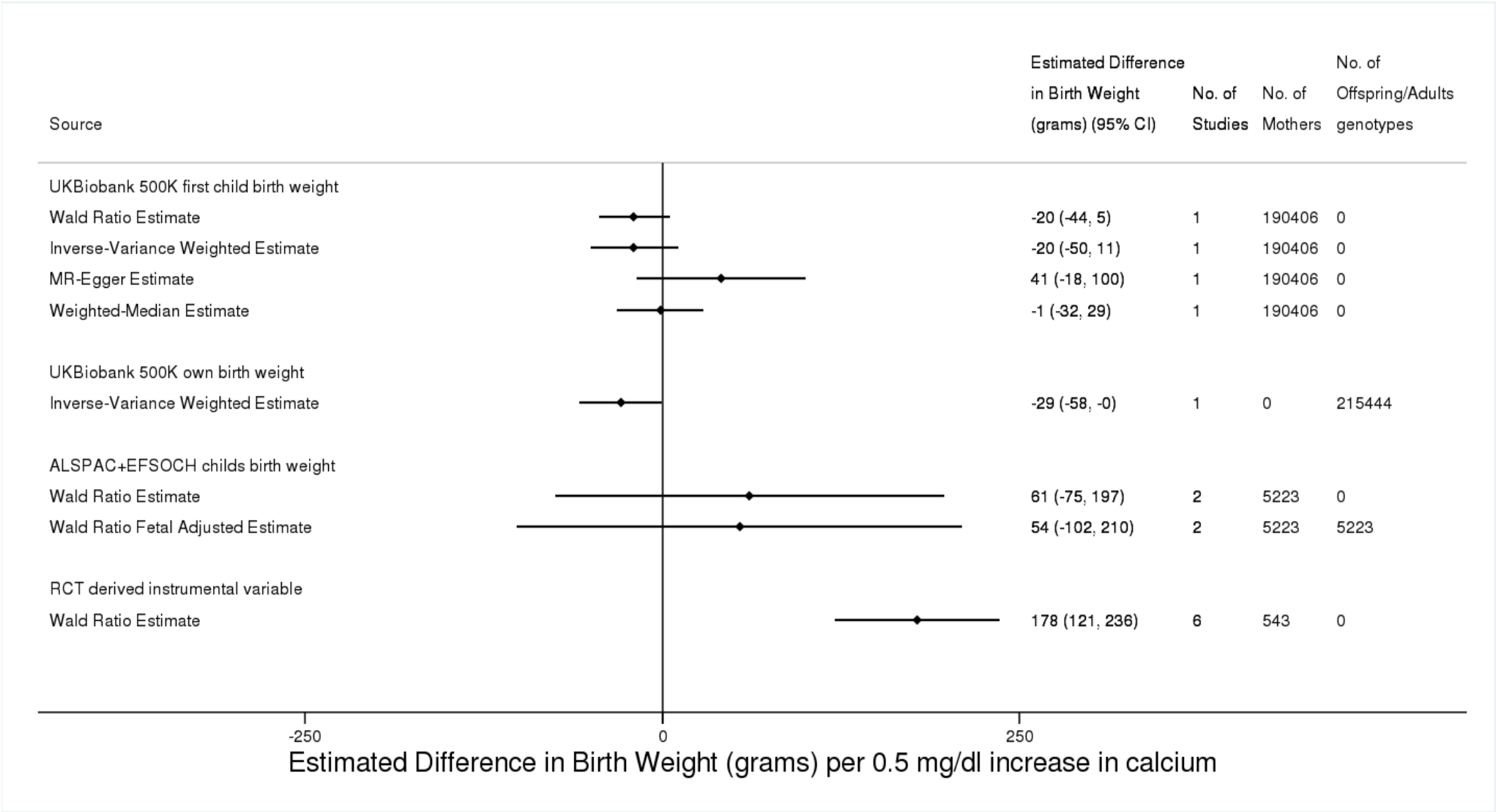
Though there were only a small number of genetic instruments, and thus limited ability to test potential biases from horizontal pleiotropy and weak instruments, there was no strong evidence that 25(OH)D synthesis or metabolism had an effect on BW (**S6 Fig**).

Mendelian randomisation and instrumental variable analyses applied to RCTs give conflicting results on the effect of calcium on birth weight

Our main two-sample MR results provide no strong evidence of an effect of maternal circulating calcium on offspring BW, with consistent null findings across the Wald ratio, IVW and weighted median results using UKB. However, there was evidence of horizontal pleiotropy, with the MR-Egger intercept being -2g (95% CI, -4 to -0.5g, p=0.025) and the effect estimate suggesting a possible modest effect of 41g (95% CI, -18 to 100g, p=0.132) per 1SD increase in maternal circulating calcium, compared to the weak inverse effect in the IVW and Wald ratio estimates of -20g (95% CI, -44 to 5g, p=0.116) (**Fig 3**). There was also evidence for heterogeneity between the Wald ratio estimates for each SNP (I^2 =62.4%), and removing one of the SNPs (rs1801725) strengthened the inverse effect estimate to -61g (95% CI, -99 to -23g, p=0.002) per 1SD increase in maternal circulating calcium, with removal of

the other six SNPs resulting in estimates close to the main result and consistent with each other (**S7 Fig**). The Wald ratio for only the rs1801725 SNP was 10g (95% CI, -22 to 41g, p=0.558) per 1 SD increase in maternal circulating calcium. There was no evidence that fetal genotype influenced the results, as results were similar with or without adjustment for it, and the IVW analyses for adults own BW gave a similar result to the IVW analyses for first child's BW (**Fig 3)**. The RCT instrumental variable effect on calcium suggested a strong positive effect that was inconsistent with any of the MR estimates (including MR-Egger): 178g (95% CI, 121 to 236g, p=1.43 x 10^{-9}) change in BW per 1SD increase in maternal circulating calcium **(Fig 3)**. There was no strong evidence for heterogeneity between the RCT instrumental variable results (I^2= 0.0%), but the main result and leave one out analyses had wide confidence intervals **(S8 Fig)**. Given the marked difference in the instrumental variable results in RCTs compared with any of the MR analyses we (*post hoc*) looked for potential bias within the specific RCTs that we were able to include in our one-sample instrumental variable analyses of calcium on BW. Of the six included RCTs, three had fewer than 100 participants (Ns = 23 to 72), and the two largest RCTs (N = 120 and 274), together with two other studies, did not use intention to treat analysis (**S9 Table**), which could result in biased effect estimates. The second largest study (N = 120) was also judged by the authors of the systematic review to have high risk of bias (or it was noted that there was insufficient information to determine risk) across all of the specific domains they assessed (**S9 Table**).

Figure 3: Causative effect estimates for maternal calcium on birth weight



Validity of the 25(OH)D and calcium genetic instruments

The magnitude of the 25(OH)D SNPs association with maternal circulating 25(OH)D in ALSPAC was similar to the magnitudes reported in the source GWAS for five of the seven SNPs. There were differences for rs8018720 (which was weakly inversely, rather than positively associated with 25(OH)D in ALSPAC) and for rs117913124 (a weaker positive difference in ALSPAC when compared to source GWAS) (**Table 3)**. Despite these differences, we found no evidence of heterogeneity between the Wald ratios and leave-one-out analyses found no evidence of an outlier effect (see above).

Table 3: 25(OH)D instruments effect on exposure in GWAS Source and ALSPAC

SNP	Nearby Genes	Effect Allele	Effect Allele Frequency (GWAS reported)	GWAS	No of Participants in GWAS	Difference in mean 25(OH)D per allele from GWAS, in natural logged (log) nmol/l (95% CI)	No of Participants in ALSPAC	Difference in mean 25(OH)D per allele from ALSPAC, in log nmol/l (95% CI)
rs3755967	*GC*	C	0.72	Jiang et al[19]	79,366	0.089 (0.084 to 0.094)	4,874	0.069 (0.048 to 0.091)
rs117913124	*CYP2R1*	G	0.975	Manousaki et al[24]	42,274	0.21 (0.19 to 0.23)*	4,874	0.079 (0.018 to 0.139)
rs10741657	*CYP2R1*	A	0.4	Jiang et al[19]	79,366	0.031 (0.027 to 0.035)	4,874	0.01 (-0.010 to 0.029)
rs12785878	*DHCR7*	T	0.75	Jiang et al[19]	79,366	0.036 (0.032 to 0.04)	4,874	0.049 (0.026 to 0.071)
rs10745742	*AMDHD1*	T	0.41	Jiang et al[19]	79,366	0.017 (0.013 to 0.021)	4,874	0.029 (0.009 to 0.049)
rs8018720	*SEC23A*	G	0.27	Jiang et al[19]	79,366	0.017 (0.012 to 0.022)	4,874	-0.037 (-0.062 to -0.011)
rs17216707	*CYP24A1*	T	0.79	Jiang et al[19]	79,366	0.026 (0.021 to 0.031)	4,874	0.020 (-0.005 to 0.046)

*This result differs from that reported by Manousaki et al. Their result was in units of standard deviations of 25(OH)D in log nmol/L. We converted that value to natural logged nmol/l by estimating the value of a standard deviation of log transformed nmol/L of 25(OH)D in the ALSPAC study.

Neither the 25(OH)D or the calcium genetic instrumental variable weighted allele scores were associated with observed confounders in the UKB, ALSPAC or EFSOCH, with the exceptions of maternal height, education and TDI. The score for 25(OH)D was negatively associated with height in the UKB, positively associated with height and negatively associated with TDI in EFSOCH and showed no association in ALSPAC; the score for calcium was negatively associated with educational level and positively associated with TDI in UKB (**S10 Table)**. Given these findings, and on advice from one of the reviewers, we undertook further analyses, using multivariable MR[47], to explore whether maternal height might have masked a positive effect of 25(OH)D on BW (maternal height could be a masking confounder in these results as it is inversely associated with 25(OH)D but positively relates to BW) and whether maternal education confounded the MR effect estimate of maternal calcium on BW (further details available in **S5 Text)**. Results from the multivariable IVW MR analyses were consistent with those from the unadjusted IVW MR analyses, for both the full adjustment of height for 25(OH)D-BW results and the partial maternal education calcium-BW results (**S11 Table**).

Discussion

This study triangulated two approaches (two-sample MR and instrumental variables applied to RCTs) assessing the effects of maternal circulating 25(OH)D and calcium on BW. Across the main, and all sensitivity, MR analyses we found no evidence that maternal 25(OH)D has an important effect on BW but applying instrumental variable analyses to RCTs there was evidence of a weak positive effect. Findings for maternal circulating calcium were inconsistent across methods and sensitivity analyses, making it difficult for us to conclude from the data used here what the effect of maternal circulating calcium on BW is.

Whilst the instrumental variable applied to RCTs analyses for 25(OH)Ds effect suggested a weak positive effect of 5.94g (95% CI, 2.15 to 9.73, p=0.002) higher birth weight per 10% increase in 25(OH)D), this might be exaggerated by limitations of the original RCTs (**S8 Table** and recent systematic review[16]) and is so small that it is unlikely to be of clinical or public health importance. Our MR analyses are largely in European origin populations whereas the RCTs were predominantly in South Asian or Middle Eastern populations and several were also in those with low 25(OH)D levels at the start of pregnancy. These population differences may also have contributed to differences between the two approaches, though it is notable that the RCTs suggest little evidence of an effect on birth weight even in these populations at high risk of vitamin D insufficiency. One of the proposed mechanisms underlying the hypothesis that maternal circulating 25(OH)D results in higher BW is by increasing offspring bone mineral density (BMD). However, a recent large, well conducted RCT found no evidence of an effect of maternal vitamin D_3 supplementation on offspring neonatal BMD (assessed within 2 weeks of birth) [48]. When combined with our own findings there is, therefore, little evidence to suggest that maternal circulating 25(OH)D influences BW via increases in BMD. A recent MR study suggested no evidence of circulating 25(OH)D affecting the risk of pre-eclampsia[49], which is consistent with results

of a meta-analysis of RCTs of vitamin D supplementation on pre-eclampsia.[16] That same meta-analysis found no strong evidence that randomisation to vitamin D supplementation influenced gestational diabetes risk, when restricting analyses to trials with least risk of bias RCTs[16]. As pre-eclampsia and gestational diabetes affect BW [6], these findings are consistent with our results suggesting maternal 25(OH)D does not affect BW.

It is possible that low circulating 25(OH)D levels are not a suitable indicator of vitamin D deficiency, as 1,25(OH)2D is the biologically active form of vitamin D and remains within reference limits even when circulating 25(OH)D levels are low, suggesting valid genetic and RCT instruments associated with 1,25(OH)2D levels could give different results. However, we are not aware of such instruments and circulating 25(OH)D has a longer half-life in the body, making it a more stable measurement and possibly a better indicator of long term vitamin D exposure[50]. Furthermore, previous observational studies that underpin the hypothesis that vitamin D levels are importantly related to BW, and a large number of other health outcomes, have used 25(OH)D as the marker of exposure.

Although the main MR analyses did not support a causal effect of maternal calcium on offspring BW, sensitivity (MR-Egger and leave-one-out analyses) were not completely consistent with the main MR effect estimates and the instrumental variable analyses applied to RCTs were markedly different to any of the MR results. There are a number of possible reasons why we might have found these differences. The MR estimate may be biased by masking horizontal pleiotropy, which both our MR-Egger and leave one out analyses suggest. Thus, the MR-Egger effect estimate suggests a modest positive effect with wide confidence intervals and the leave one out analyses suggests that the heterogeneity between individual SNP Wald ratios is driven by one SNP rs1801725 which when removed results in an inverse association (again suggesting masking pleiotropy). Rs1801725 is in the *CASR* gene, which codes for the calcium-sensing receptor protein and is widely express in the parathyroid gland,

kidneys and intestines and regulated blood levels of calcium [51], whereas the other calcium SNPs are near to loci that are not known to have such a direct effect on calcium levels and may therefore be more prone to pleiotropy. The leave one out analyses, together with the Wald ratio result for *CASR* rs1801725 alone and the MR-Egger result all suggest that our main IVW results might be biased by masking horizontal pleiotropy and that there is a modest positive causal effect of calcium on BW. Possible sources of masking pleiotropy could be alterations to glucose levels, as one of the genetic instruments (rs780094) is associated with fasting glucose which is known to influence BW[52]. However, whilst these sensitivity analyses suggest a possible modest positive causal effect of maternal circulating calcium on BW, they have limited statistical power and therefore they are imprecisely estimated with wide confidence intervals. Importantly, whilst we were able to show most of the 25(OH)D SNPs related to pregnancy 25(OH)D, we were not able to assess this for maternal gestational calcium. If the calcium SNPs have weaker associations with circulating calcium levels in pregnancy (than they do in the GWAS of men [25]) this might bias towards the null any real effect of maternal pregnancy calcium on BW in our analyses. During pregnancy maternal circulating levels of calcium increase in order to support healthy skeletal development. This is achieved through increasing absorption from the intestines and through bone resorption in the mothers, but this process is, at least partially, under the control of the fetus, which increases the secretion of parathyroid hormone-related protein in response to low fetal plasma calcium [53]. Thus, fetal genetic variants might be important genetic instruments for MR analyses of maternal circulating gestational calcium's effect on fetal skeletal development and hence BW.

With respect to the difference between MR and instrumental variable analyses in RCT results, we might expect to see weaker effects from instrumental variable analyses in RCTs than from MR analyses as the former only tests differences in exposure from the time of

randomisation, whereas MR tests differences over most of life (and hence across all pregnancy trimesters). On-the-other-hand, for in utero exposures timing rather than duration of exposure might be important[20]. Though the fetus acquires all of its calcium from the mother throughout pregnancy, some evidence suggests that 80% of calcium in a neonates bones are absorbed during the third trimester (possibly due to an increase in maternal 1,25-dihydroxyvitamin D levels in the third trimester) [54], which would mean that supplementation starting earlier in pregnancy and continuing through to delivery might be necessary to ensure adequate levels for bone development throughout the third trimester (and hence an effect on BW). Five of the six RCTs that we included in our analyses would fit with this, as supplementation began at or before 23 weeks of gestation and continued until delivery (the one study which start supplementation in the third trimester 28 to 31 weeks only included 32 participants). Overall, we might then expect the RCT and MR results to be similar. However, the results from the instrumental variable analyses in RCTs might be biased due to the inclusion of only six calcium supplementation RCTs, with the two largest ones having important sources of bias (**S9 Table**). Furthermore, whilst our MR analyses are largely in European origin populations, the RCTs were predominantly in non-European populations, in particular Latin American populations, which may explain some of the differences.

Study Strengths and Limitations

To the best of our knowledge, this is the first study to compare results from MR and instrumental variable analyses applied to RCTs to investigate the effect of maternal circulating calcium levels on offspring BW. The GWAS that we used for genetic instrument-exposure associations (sample 1) in our two sample MR did not provide the percentage variation in calcium that all seven SNPs explained. However, based on a previous study that provided the R^2 for one of these seven SNPs (rs1801725)[55] we know that our instrument

with all seven SNPs will explain at least 2% of the variation in circulating calcium. We have considerably increased the sample size of our previous MR study of the effect of maternal circulating 25(OH)D on offspring BW[6]. These additions have increased the strength of our instrument, with the R^2 in our study suggesting the genetic variants explained ~3% of variation in circulating 25(OH)D compared with <1% of the variation explained by the two SNPs used in the previous MR study[6].We have explored the validity of our genetic instrumental variables using multiple sensitivity analyses and compared those results with instrumental variable analyses applied to RCTs, in which summary results from the RCTs were extracted independently by two people and checked for consistency with a third.

Key potential limitations include the low response to UKB and maternal-report of first child's BW many years after their birth. Recent research suggests that a highly select cohort (as in the case of UKB with a 5% response [56]) can result in selection bias in genetic or MR analyses [57,58]. Self-report of BW and the rounding to 1 pound (~0.454kg) may have introduced error in the BW measure in UKB, but this would be random with respect to genotype and expected to bias results towards the null. The somewhat lower mean BW in UKB participants than ALSPAC/EFSOCH is likely to reflect secular trends of increasing birth size over time the fact that they reported only the weight of first-born children, whereas in ALSPAC/EFSOCH index children are from any pregnancy. The ALSPAC/EFSOCH MR estimates were stronger than the UKB estimates, though with very wide confidence intervals reflecting their smaller sample size and there is no evidence that the results were statistically inconsistent with those from UKB. These two cohorts were used in sensitivity analyses to explore whether the main MR analyses might be biased by a path from maternal genotype via fetal genotype to their BW. We have shown this is not the case. However, adjustment for fetal genotype could introduce a spurious association between maternal and paternal genotype[44], and if fathers circulating 25(OH)D or calcium influenced offspring BW this

could bias our results. As it is unlikely that fathers circulating levels of 25(OH)D or calcium could directly influence fetal growth and offspring BW we feel our offspring adjusted results are unlikely to be biased. This is also supported by the lack of any effect of own 25(OH)D or calcium on own BW in MR sensitivity analyses in UKB.

In UKB the 25(OH)D genetic risk score was negatively associated with height, and as it is possible that greater maternal height results in greater BW, independently of the offspring genotype[59], this could mask any true positive effect of 25(OH)D on BW. We also found that the calcium genetic instrument was associated with markers of socioeconomic position (area deprivation and education), which could confound the MR estimated effect of calcium on BW. However, results of multivariable MR adjusting the maternal 25(OH)D effect on BW for maternal height and the maternal calcium effect on BW (partially) for maternal education were consistent with the unadjusted results, suggesting that our results were not importantly confounded.

Horizontal pleiotropy is a key source of bias for MR studies. For the effects of maternal circulating 25(OH)D on BW the consistency of results across all of our MR analyses suggest that has not been a key source of bias ~~in~~ for that exposure. We did find evidence of potential horizontal pleiotropy with maternal calcium effects and as discussed above, and acknowledge that further studies are required to explore that effect. We have assumed that genetic variants that relate to 25(OH)D/calcium in (non-pregnant) women and men have the same magnitudes of association with these exposures in pregnant women. We were unable to test this for calcium but did show that five out of the seven variants related similarly to circulating 25(OH)D during pregnancy as in the GWAS. Removal of either of the two that did not associate as strongly to 25(OH)D in pregnancy in our leave-one-out analyses were consistent with the main results and all other leave-one-out analyses. Although we aimed to use randomisation to use supplementation with vitamin D or calcium as an instrumental variable

for RCTs, some studies did not provide results from intention to treat analyses. Therefore, results from these studies could be biased by non-compliance. The aim of triangulation is to compare results from different methods that have different key sources of bias. Whilst both approaches that we have used here rely on instrumental variables and have the same underlying assumptions, the sources of violation of these assumptions differ between MR and instrumental variables applied to RCTs. For MR, horizontal pleiotropy is the key source of bias, whereas for instrumental variables applied to RCT, it is lack of concealed random allocation (which can introduce bias) and not using intention to treat analysis (which can produce a path from the instrument to outcome and/or introduce confounding). The fact that there was no marked difference between our MR and RCT analyses strengthen confidence in the findings for 25(OH)D. However, as discussed above, the different results for calcium need further exploration. In both our two sample MR and instrumental variable analyses applied to RCTs we have used summary data and are unable to explore possible non-linear effects, though observational studies do not suggest that these are present. Similarly, we cannot explore effect modification, but we are not aware of any evidence that the effects that we have looked at do differ by other characteristics. Our MR analyses are in white European populations and may not generalise to other groups, such as those with different levels of skin pigmentation and exposures to sunshine. However, most of the RCTs for vitamin D and calcium supplementation are in non-Europeans. In particular most vitamin D RCTs were in Middle Eastern or South Asian populations and a recent RCT in Bangladesh, which was not included in the systematic review that we used in this paper for the RCT analyses, found that vitamin D supplementation during pregnancy made no difference to BW [60]. Thus even in a low-income, dark skinned population, maternal circulating 25(OH)D may not be an important factor in BW.

In conclusion, our results suggest that maternal circulating 25(OH)D does not have a clinically important effect on BW., so pregnancy supplementation with vitamin D is unlikely to affect mean BW. Higher maternal circulating calcium may increase BW, but further research is required to clarify this, including larger samples for undertaking pleiotropy 'adjusted' MR analyses and larger better conducted RCTs of calcium supplementation.

Acknowledgements

This research has been conducted using the UK Biobank Resource under application number 7036. We would like to thank the participants and researchers from the UK Biobank who contributed or collected data and the families that took part in EFSOCH. We are extremely grateful to all of the families who took part in ALSPAC, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We are grateful to the Genetics of Complex traits team at the University of Exeter, for their assistance in learning the methods and navigating the study data. We are particularly grateful to Hanieh Yaghootkar, for reading a Persian language paper for us and providing the key data.

References

1. Iliodromiti S, Mackay DF, Smith GC, Pell JP, Sattar N, Lawlor DA, et al. Customised and Noncustomised Birth Weight Centiles and Prediction of Stillbirth and Infant Mortality and Morbidity: A Cohort Study of 979,912 Term Singleton Pregnancies in Scotland. PLoS Med. 2017;14(1):e1002228. Epub 2017/02/01. doi: 10.1371/journal.pmed.1002228.
2. Wang SF, Shu L, Sheng J, Mu M, Wang S, Tao XY, et al. Birth weight and risk of coronary heart disease in adults: a meta-analysis of prospective cohort studies. Journal of developmental origins of health and disease. 2014;5(6):408-19. Epub 2014/09/30. doi: 10.1017/s2040174414000440.
3. Whincup PH, Kaye SJ, Owen CG, Huxley R, Cook DG, Anazawa S, et al. Birth weight and risk of type 2 diabetes: a systematic review. Jama. 2008;300(24):2886-97. Epub 2008/12/26. doi: 10.1001/jama.2008.886.
4. Yang S, Lynch J, Susser ES, Lawlor DA. Birth weight and cognitive ability in childhood among siblings and nonsiblings. Pediatrics. 2008;122(2):e350-8. Epub 2008/08/05. doi: 10.1542/peds.2007-3851.
5. Ross JA. High birthweight and cancer: evidence and implications. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2006;15(1):1-2. Epub 2006/01/26. doi: 10.1158/1055-9965.epi-05-0923.
6. Tyrrell J, Richmond RC, Palmer TM, et al. Genetic evidence for causal relationships between maternal obesity-related traits and birth weight. JAMA. 2016;315(11):1129-40. doi: 10.1001/jama.2016.1975.
7. Brand JS, West J, Tuffnell D, Bird PK, Wright J, Tilling K, et al. Gestational diabetes and ultrasound-assessed fetal growth in South Asian and White European women: findings from a prospective pregnancy cohort. BMC Medicine. 2018;16(1):203. doi: 10.1186/s12916-018-1191-7.
8. Specker B. Vitamin D requirements during pregnancy. The American Journal of Clinical Nutrition. 2004;80(6):1740S-7S. doi: 10.1093/ajcn/80.6.1740S.

9. Bergel E, Belizán JM. Commentary: Maternal calcium intake and offspring cardiovascular risk factors. International Journal of Epidemiology. 2004;33(6):1309-10. doi: 10.1093/ije/dyh362.
10. Tian Y, Holzman C, Siega-Riz AM, Williams MA, Dole N, Enquobahrie DA, et al. Maternal Serum 25-Hydroxyvitamin D Concentrations during Pregnancy and Infant Birthweight for Gestational Age: A Three-Cohort Study. Paediatric and perinatal epidemiology. 2016;30(2):124-33. doi: 10.1111/ppe.12262.
11. Chen YH, Fu L, Hao JH, Yu Z, Zhu P, Wang H, et al. Maternal vitamin D deficiency during pregnancy elevates the risks of small for gestational age and low birth weight infants in Chinese population. The Journal of clinical endocrinology and metabolism. 2015;100(5):1912-9. Epub 2015/03/17. doi: 10.1210/jc.2014-4407.
12. Yang C-Y, Chiu H-F, Chang C-C, Wu T-N, Sung F-C. Association of Very Low Birth Weight with Calcium Levels in Drinking Water. Environmental Research. 2002;89(3):189-94. doi: https://doi.org/10.1006/enrs.2002.4369.
13. Gascoin G, Gerard M, Sallé A, Becouarn G, Rouleau S, Sentilhes L, et al. Risk of low birth weight and micronutrient deficiencies in neonates from mothers after gastric bypass: a case control study. Surgery for Obesity and Related Diseases. 2017;13(8):1384-91. doi: https://doi.org/10.1016/j.soard.2017.03.017.
14. Roth DE, Leung M, Mesfin E, Qamar H, Watterworth J, Papp E. Vitamin D supplementation during pregnancy: state of the evidence from a systematic review of randomised trials. BMJ. 2017;359:j5237. doi: 10.1136/bmj.j5237.
15. Buppasiri P, Lumbiganon P, Thinkhamrop J, Ngamjarus C, Laopaiboon M, Medley N. Calcium supplementation (other than for preventing or treating hypertension) for improving pregnancy and infant outcomes. Cochrane Database of Systematic Reviews. 2015;(2). doi: 10.1002/14651858.CD007079.pub3.
16. Roth DE, Leung M, Mesfin E, Qamar H, Watterworth J, Papp E. Vitamin D supplementation during pregnancy: state of the evidence from a systematic review of randomised trials. BMJ. 2017;359.
17. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Statistics in medicine. 2008;27(8):1133-63. Epub 2007/09/22. doi: 10.1002/sim.3034.
18. Beaumont RN, Warrington NM, Cavadino A, Tyrrell J, Nodzenski M, Horikoshi M, et al. Genome-wide association study of offspring birth weight in 86 577 women identifies five novel loci and highlights maternal genetic effects that are independent of fetal genetics. Human molecular genetics. 2018;27(4):742-56. Epub 2018/01/09. doi: 10.1093/hmg/ddx429.
19. Jiang X, O'Reilly PF, Aschard H, Hsu YH, Richards JB, Dupuis J, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nature communications. 2018;9(1):260. Epub 2018/01/19. doi: 10.1038/s41467-017-02662-2.
20. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. Int J Epidemiol. 2016;45(6):1866-86. Epub 2017/01/22. doi: 10.1093/ije/dyw314.
21. Townsend P, Phillimore P, Beattie A. Health and Deprivation: Inequality and the North. London: Routledge; 1988.
22. Bradbury KE, Young HJ, Guo W, Key TJ. Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. Journal of Nutritional Science. 2018;7:e6. doi: 10.1017/jns.2017.66.
23. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. International Journal of Epidemiology. 2016;45(3):908-15. doi: 10.1093/ije/dyw127.
24. Manousaki D, Dudding T, Haworth S, Hsu Y-H, Liu C-T, Medina-Gómez C, et al. Low-Frequency Synonymous Coding Variation in CYP2R1 Has Large Effects on Vitamin D Levels and Risk of Multiple Sclerosis. The American Journal of Human Genetics. 2017;101(2):227-38. doi: https://doi.org/10.1016/j.ajhg.2017.06.014.

25. O'Seaghdha CM, Wu H, Yang Q, Kapur K, Guessous I, Zuber AM, et al. Meta-Analysis of Genome-Wide Association Studies Identifies Six New Loci for Serum Calcium Concentrations. PLOS Genetics. 2013;9(9):e1003796. doi: 10.1371/journal.pgen.1003796.
26. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. International Journal of Epidemiology. 2013;42(1):97-110. doi: 10.1093/ije/dys066.
27. Knight B, Shields BM, Hattersley AT. The Exeter Family Study of Childhood Health (EFSOCH): study protocol and methodology. Paediatric and Perinatal Epidemiology. 2006;20(2):172-9. doi: 10.1111/j.1365-3016.2006.00701.x.
28. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203-9. doi: 10.1038/s41586-018-0579-z.
29. Lawlor DA, Wills AK, Fraser A, Sayers A, Fraser WD, Tobias JH. Association of maternal vitamin D status during pregnancy with bone-mineral content in offspring: a prospective cohort study. Lancet. 2013;381(9884):2176-83. doi: 10.1016/S0140-6736(12)62203-X.
30. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. bioRxiv. 2017. doi: 10.1101/166298.
31. UKBiobank, theAccessTeam. Important note about imputed genetics data biobank2017 [cited 2018 1st of March]. Available from: http://www.ukbiobank.ac.uk/2017/07/important-note-about-imputed-genetics-data/. (last accessed 2019 2nd of May)
32. Loh P-R, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed model analysis increases association power in large cohorts. Nature genetics. 2015;47(3):284-90. doi: 10.1038/ng.3190.
33. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort Profile: The 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. International Journal of Epidemiology. 2013;42(1):111-27. doi: 10.1093/ije/dys064.
34. Wang TJ, Zhang F, Richards JB, Kestenbaum B, van Meurs JB, Berry D, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. Lancet. 2010;376(9736):180-8. doi: 10.1016/S0140-6736(10)60588-0.
35. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics (Oxford, England). 2015;31(21):3555-7. Epub 2015/07/04. doi: 10.1093/bioinformatics/btv402.
36. Machiela MJ, Chanock SJ. LDassoc: an online tool for interactively exploring genome-wide association study results and prioritizing variants for functional investigation. Bioinformatics (Oxford, England). 2018;34(5):887-9. Epub 2017/10/03. doi: 10.1093/bioinformatics/btx561.
37. Vimaleswaran KS, Berry DJ, Lu C, Tikkanen E, Pilz S, Hiraki LT, et al. Causal Relationship between Obesity and Vitamin D Status: Bi-Directional Mendelian Randomization Analysis of Multiple Cohorts. PLOS Medicine. 2013;10(2):e1001383. doi: 10.1371/journal.pmed.1001383.
38. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. International Journal of Epidemiology. 2016;45(6):1717-26. doi: 10.1093/ije/dyx028.
39. Goldstein D. Serum Calcium. 1990. In: Clinical Methods: The History, Physical, and Laboratory Examinations [Internet]. Boston: Butterworths. 3rd. Available from: https://www.ncbi.nlm.nih.gov/books/NBK250/.(last accessed 2019 2nd of May)
40. Stephen B, Dylan SS, Simon GT. A review of instrumental variable estimators for Mendelian randomization. Statistical Methods in Medical Research. 2015;26(5):2333-55. doi: 10.1177/0962280215597579.
41. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International Journal of Epidemiology. 2015;44(2):512-25. doi: 10.1093/ije/dyv080.

42. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genetic Epidemiology. 2016;40(4):304-14. doi: 10.1002/gepi.21965.
43. Steichen T. METANINF: Stata Module to Evaluate Influence of a Single Study in Meta-analysis Estimation2001.
44. Lawlor D, Richmond R, Warrington N, McMahon G, Davey Smith G, Bowden J, 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 research. 2017;2:11-. doi: 10.12688/wellcomeopenres.10567.1.
45. Greenland S. An introduction to instrumental variables for epidemiologists. International Journal of Epidemiology. 2000;29(4):722-9. doi: 10.1093/ije/29.4.722.
46. Boggess KA, Samuel L, Schmucker BC, Waters J, Easterling TR. A randomized controlled trial of the effect of third-trimester calcium supplementation on maternal hemodynamic function. Obstetrics & Gynecology. 1997;90(2):157-61. doi: https://doi.org/10.1016/S0029-7844(97)00248-2.
47. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol. 2018:dyy262. Epub 2018/12/12. doi: 10.1093/ije/dyy262.
48. Cooper C, Harvey NC, Bishop NJ, Kennedy S, Papageorghiou AT, Schoenmakers I, et al. Maternal gestational vitamin D supplementation and offspring bone health (MAVIDOS): a multicentre, double-blind, randomised placebo-controlled trial. The Lancet Diabetes & Endocrinology. 2016;4(5):393-402. doi: http://dx.doi.org/10.1016/S2213-8587(16)00044-9.
49. Magnus MC, Miliku K, Bauer A, Engel SM, Felix JF, Jaddoe VWV, et al. Vitamin D and risk of pregnancy related hypertensive disorders: mendelian randomisation study. BMJ. 2018;361:k2167. doi: 10.1136/bmj.k2167.
50. Lips P. Relative Value of 25(OH)D and 1,25(OH)2D Measurements. Journal of Bone and Mineral Research. 2007;22(11):1668-71. doi: 10.1359/jbmr.070716.
51. Pu F, Chen N, Xue S. Calcium intake, calcium homeostasis and health. Food Science and Human Wellness. 2016;5(1):8-16. doi: https://doi.org/10.1016/j.fshw.2016.01.001.
52. Bi M, Kao WHL, Boerwinkle E, Hoogeveen RC, Rasmussen-Torvik LJ, Astor BC, et al. Association of rs780094 in GCKR with metabolic traits and incident diabetes and cardiovascular disease: the ARIC Study. PloS one. 2010;5(7):e11690-e. doi: 10.1371/journal.pone.0011690.
53. Goodfellow L, Cooper C, Harvey N. Regulation of Placental Calcium Transport and Offspring Bone Health. Frontiers in Endocrinology. 2011;2(3). doi: 10.3389/fendo.2011.00003.
54. CS. K. Calcium Metabolism during Pregnancy and Lactation. 2000. In: Endotext [Internet]. South Dartmouth (MA): MDText.com, Inc. Available from: https://www.ncbi.nlm.nih.gov/books/NBK279173/.(last accessed 2019 2nd of May)
55. Kapur K, Johnson T, Beckmann ND, Sehmi J, Tanaka T, Kutalik Z, et al. Genome-wide meta-analysis for serum calcium identifies significantly associated SNPs near the calcium-sensing receptor (CASR) gene. PLoS Genet. 2010;6(7):e1001035. Epub 2010/07/28. doi: 10.1371/journal.pgen.1001035.
56. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. American journal of epidemiology. 2017;186(9):1026-34. Epub 2017/06/24. doi: 10.1093/aje/kwx246.
57. Munafo MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47(1):226-35. Epub 2017/10/19. doi: 10.1093/ije/dyx206.
58. Sallis H, Taylor AE, Munafò MR, Stergiakouli E, Euesden J, Davies NM, et al. Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. International Journal of Epidemiology. 2018;47(4):1207-16. doi: 10.1093/ije/dyy060.

59. Warrington NM, Beaumont RN, Horikoshi M, Day FR, Helgeland Ø, Laurin C, et al. Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. Nature genetics. 2019. doi: 10.1038/s41588-019-0403-1.
60. Roth DE, Morris SK, Zlotkin S, Gernand AD, Ahmed T, Shanta SS, et al. Vitamin D Supplementation in Pregnancy and Lactation and Infant Growth. New England Journal of Medicine. 2018;379(6):535-46. doi: 10.1056/NEJMoa1800927.

Supporting Information

S1 Checklist

S1 Text: Study Descriptions

S2 Text: Selecting of participants of White European ancestry

S3 Text: Selecting participants for own birth weight analyses in UK Biobank

S4 Text: Measuring 25(OH)D during gestation in mothers in the ALSPAC cohort

S5 Text: Sensitivity analysis to explore additional sources of invalid instruments

S1 Table: Characteristics of the genome-wide association studies of 25(OH)D

S2 Table: Details of SNPs used in our Mendelian Randomisation analyses

S3 Table: Studies used to calculate the RCT instrumental variable effect of 25(OH)D on birth weight

S4 Table: Studies used to calculate the RCT instrumental variable effect of calcium on birth weight

S5 Table: SNP effects on first child birth weight in all studies

S6 Table: SNP effects on own birth weight in UK Biobank (N=215,444)

S7 Table: SNP effects on fetal adjusted birth weight in ALSPAC and EFSOCH

S8 Table: Risk of Bias in studies included in IV of RCT analyses of Vitamin D supplementation

S9 Table: Risk of Bias in studies included in IV of RCT analyses of Calcium supplementation

S10 Table: Associations between weighted-allele-scores and potential confounders

S11 Table: Multivariable MR for 25(OH)D and calcium effect on birth weight in UK Biobank (adjusting for height effects)

S1 Fig: Flow diagram of participant inclusion for ALSPAC and EFSOCH

S2 Fig: Flow diagram for participant inclusion in UK Biobank

S3 Fig: Flow diagram for inclusion of trials in the instrumental variables applied to RCTs

S4 Fig: Leave-One-Out Analysis for effect of maternal gestational circulating 25(OH)D on birth weight Mendelian randomisation Wald ratio estimate in UK Biobank

S5 Fig: Leave-One-Out Analysis for effect of 25(OH)D on birth weight RCT instrumental variable Wald ratio estimate

S6 Fig: Mendelian randomisation effect estimates for maternal 25(OH)D synthesis and metabolism on birth weight in UK Biobank

S7 Fig: Leave-One-Out Analysis for effect of maternal gestational circulating calcium on birth weight Mendelian randomisation Wald ratio estimate in UK Biobank

S8 Fig: Leave-One-Out Analysis for effect of maternal gestational circulating calcium RCT instrumental variable Wald ratio estimate