

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Maternal caffeine consumption during pregnancy and offspring cord blood DNA methylation: a meta-analysis of epigenome-wide association studies

Laura Schellhas^{1,2,3}, Giulietta S Monasso^{4,5}, Janine F Felix^{4,5}, Vincent WV Jaddoe^{4,5}, Peiyuan Huang², Sílvia Fernández-Barrés^{6,9}, Martine Vrijheid^{6,7,8}, Giancarlo Pesce¹², Isabella Annesi-Maesano^{10,11}, Christian M Page¹³, Anne-Lise Brantsæter¹⁴, Mona Bekkhus¹⁶, Siri E Håberg¹⁵, Stephanie J London¹⁷, Marcus R Munafò^{1,2,20}, Luisa Zuccolo^{2,18,19}, Gemma C Sharp^{2,18,21}

1. School of Psychological Science, University of Bristol, Bristol, UK
2. MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK
3. Institute for Sex Research and Forensic Psychiatry, University Medical Center Hamburg-Eppendorf
4. The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
5. Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
6. Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain
7. Universitat Pompeu Fabra (UPF), Barcelona, Spain
8. CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
9. Agència de Salut Pública de Barcelona, Pl. Lesseps 1, 08023, Barcelona, Spain
10. Institute Desbrest of Epidemiology and Public Health, INSERM and Montpellier University
11. Department of Allergic and Respiratory Diseases, Montpellier University Hospital, Montpellier, France
12. INSERM UMR-S 1136, Team of Epidemiology of Allergic and Respiratory Diseases (EPAR), Institute Pierre Louis of Epidemiology and Public Health (IPLESP), Sorbonne University, Paris, France.
13. Department of Physical Health and Aging, Division for Mental and Physical Health, Norwegian Institute of Public Health, Oslo, Norway
14. Department of Food Safety, Division of Climate and Environmental Health, Norwegian Institute of Public Health, Oslo, Norway
15. Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway
16. Department of Psychology, Promenta Research Centre, University of Oslo, Oslo, Norway
17. Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709 USA
18. Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK
19. Health Data Science Centre, Fondazione Human Technopole, Milan, Italy
20. NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol
21. School of Psychology, University of Exeter, Exeter, UK

Conflict of Interest: None declared.

E-mail: Laura.Schellhas@bristol.ac.uk, l.schellhas@uke.de

Abstract word count: 255

Main text word count: 3,654

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Abstract

Background: Women are advised to limit caffeine consumption during pregnancy. To uncover the potential epigenetic effects of intrauterine caffeine exposure, we investigated the association of maternal caffeine consumption during pregnancy with genome-wide DNA methylation in cord blood.

Methods: We meta-analysed results from epigenome-wide methylation studies at individual probes and in differentially methylated region (DMR) analysis across 6 European pregnancy and birth cohorts (ALSPAC, BiB, MoBa, Generation R, INMA, EDEN; total n = 3742). Methylation was assessed with Illumina Infinium 450k or EPIC arrays. Maternal caffeine consumption (mg/day) from coffee, tea and cola was derived from questionnaires between weeks 12 - 22 of pregnancy. We investigated associations of methylation with overall and beverage-specific caffeine intake in models adjusted for maternal education, age, BMI, smoking during pregnancy, parity, cord-blood cell proportions and 20 surrogate variables.

Results: One CpG site (cg19370043, nearest gene *PRRX1*) was associated with maternal caffeine consumption after FDR adjustment for multiple testing and one CpG sites (cg14591243, nearest gene *STAG1*) was associated with maternal cola consumption. We found evidence for 12-22 DMRs for each of the caffeine models but little overlap between DMRs identified for specific caffeinated beverages.

Conclusions: In models adjusted for maternal smoking and other potential confounders, we found little evidence to support an intrauterine effect of caffeine on offspring DNA methylation. Comparing associations across different sources of caffeine provided no evidence for caffeine being the causal agent. It is possible that our study did not have adequate statistical power to detect very small associations between maternal caffeine and offspring DNA methylation.

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Keywords: *caffeine, pregnancy, offspring health, epigenetics, DNA methylation,*

PACE

Key messages:

- This large-scale meta-analysis of epigenomewide association studies across six European cohorts does not support an intrauterine effect of caffeine on offspring cord blood DNA methylation.
- Lack of overlap between associations with different caffeinated drinks suggest that any (weak) associations were driven by diverse confounding structures of different caffeinated drinks, rather than caffeine per se.
- More research is needed to understand the biological mechanisms driving potential effects of caffeine on offspring health.

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Introduction

The effects of intrauterine caffeine exposure on offspring health outcomes are not well understood. And as a result, pregnancy guidelines for caffeine use vary during pregnancy (1). Throughout pregnancy, the metabolic rate of caffeine gradually decreases. In the second and third trimester of pregnancy the half-life of caffeine can be up until 18 hours, which is about four times longer than outside of pregnancy (2). Due to pregnancy-related changes to the caffeine metabolism and the potential of caffeine readily crossing the placenta barrier, the European Food Safety Authority (EFSA) states “...unborn children to be the most vulnerable group for adverse effects of caffeine among the general population” (3). Current pregnancy guidelines for caffeine consumption across Asia, Oceania, Europe, Latin America, and North America recommend avoiding or limiting caffeine intake during pregnancy (1). These guidelines are based evidence from observational study designs, with results likely affected by selection, measurement, and confounding biases and therefore cannot differentiate confounded from causal effects (4). This is especially problematic as caffeine consumption is strongly culturally bound, and the sources of caffeine may vary between different countries and be differentially socially patterned (1,5–7).

Identification of a molecular pathway that could explain how prenatal exposure to caffeine potentially increases the risk for adverse offspring health outcomes would add to the evidence base and strengthen the causal interpretation of these associations.

DNA methylation in particular has been proposed as the mediating mechanism between prenatal exposures and offspring health outcomes (8). In this epigenetic modification, a methyl group is added to a cytosine-phosphate-guanine (CpG) site in the genome (9). A recent epigenome-wide association study (EWAS) meta-analysis of adult populations found coffee consumption to be associated with own peripheral blood DNA methylation at 11 CpG sites (10). No published studies have examined the role of prenatal exposure

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

to caffeine in humans, but there is evidence from animal studies that prenatal caffeine consumption can lead to DNA methylation changes and offspring risk of disease (11–14).

In this study, we explored the association between maternal caffeine consumption (from tea, coffee, and cola) during pregnancy and offspring cord-blood DNA methylation across the genome, by meta-analysing results from six international pregnancy and birth cohorts from the Pregnancy and Childhood Epigenetics (PACE) consortium (8).

Methods

Meta-analysis of epigenome-wide association studies (EWAS)

Participating cohorts

The EWAS meta-analysis included six independent prospective pregnancy and birth cohorts from the PACE consortium (8), that had data available on cord blood DNA methylation and maternal caffeine consumption during pregnancy. The total sample (N = 3,731) included two UK based cohorts (ALSPAC (15,16) and BiB)(17,18), one Dutch (Generation R) (19,20), one Norwegian (MoBa1) (21), one Spanish (INMA) (22), and one French (EDEN) (23) cohort. Recruitment periods varied by cohort and took place between the beginning of the 1990s (ALSPAC) and 2008 (MoBa1). Ethical approval was obtained by local ethics committees and informed consent for the use of data was obtained from all participants. More details about the individual cohorts can be found in the supplementary information.

Measurement of maternal caffeine intake during pregnancy

Assessment of maternal caffeine consumption varied by cohort and is described in more detail in Supplementary Information. Generally, mothers self-reported the number of cups they consumed of caffeinated coffee, tea, and cola in questionnaires between week 12 to 22 of pregnancy. All cohorts used Food Frequency Questionnaires (FFQ) (24)

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

except for Generation R, which also did not have information on caffeinated cola consumption available. Cups per day were transformed to milligrams of caffeine per day (mg/day), based on the assumption that one standard sized cup of coffee contains 57 mg, one cup of tea contains 27 mg, and one cup of cola contains 20 mg of caffeine (25). A continuous total caffeine score was calculated by summing the caffeine content from each caffeinated drink in mg/day (supplementary information). In addition to the continuous score, we investigated whether any caffeine exposure (regardless of the amount of caffeine) during pregnancy might have an effect on offspring DNA methylation, and thus dichotomised total caffeine into 0 mg/day = none, and > 0 mg/day = any.

Measurement of DNA methylation

Cohorts assessed cord blood DNA methylation data individually, using their own laboratory methods, quality control, and normalisation. DNA methylation data was sampled using the Illumina Infinium® HumanMethylation450 (486,425 probes), except for BiB, which used the Illumina EPIC BeadChip array (866,553 probes). Probes on single nucleotide polymorphisms (SNPs), cross-hybridizing probes (26), and probes on the sex chromosomes were excluded. In the final meta-analysis, only probes that were available in both arrays (maximum 364,678) were included. Methylation was measured using normalised beta values ranging from 0 to 1, representing 0 to 100% methylation.

Covariates

To adjust for variation in DNA methylation driven by cell composition, models were adjusted for cell proportions estimated using the Houseman method with a cord blood reference panel (27,28). Offspring sex was used to conduct sex-stratified sensitivity analyses because sex-specific DNA methylation differences can still be observed even when restricting analyses to autosomes and removing probes that are cross-reactive

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

with sex chromosomes (29). Another sensitivity analysis was conducted where the unstratified models were additionally adjusted for gestational age at birth. Gestational age could be a mediating factor as it is robustly associated with DNA methylation (30,31) and there is some evidence that it can be associated with prenatal caffeine exposures (32,33). To avoid introducing collider bias, we adjusted for gestational age in separate models instead of the main models (34). To adjust for possible technical variation, all cohorts generated 20 surrogate variables and included them in models - as is standard practice in the field (35).

Each model contained the following covariates (35,36): An ordinal measure representing maternal education as a proxy for socio-economic position, maternal age in years, maternal BMI (kg/m^2), a binary measure of maternal smoking during pregnancy (e.g., in ALSPAC: 0 = no smoking or giving up smoking during the first trimester, 1 = smoking after the 1st trimester), and a binary assessment of parity (1 = one or more previous children, 0 = no previous children). See supplementary information for the classifications of covariates in each cohort.

Statistical analyses

Cohort-specific statistical analyses

Probe-level analysis

The analysis plan and R-script is available on GitHub

(https://github.com/ammegandchips/Prenatal_Caffeine). Cohorts were asked to

exclude multiple pregnancies (e.g., twins) and siblings so that each mother was only

included once in the dataset. If cohorts included more than one major ethnic group,

they were asked to run the EWAS analysis separately for each group. The EWAS R script

included: a function to remove probes classified as an outlier according to the Tukey

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

method of outlier removal (values < -1.5 *interquartile range and values > 1.5 *interquartile range)(37), a function to generate surrogate variables using the R package SVA (38), and a function to run an EWAS of each model using the R package Limma (39). In a second sensitivity analysis, the binary (any vs. none) and total continuous unstratified models were additionally adjusted for gestational age at birth. In a separate sensitivity analysis without gestational age adjustment, the binary (any vs. none) and total continuous caffeine models were stratified by offspring's sex. For quality assurance, an independent shadow meta-analysis was conducted by a co-author of the University of Bristol (Peiyuan Huang). Prior to meta-analysing summary results from each cohort, quality checks were conducted to ensure that the EWAS were properly conducted and there were no problems with the data, in line with standard practice in the field (35,40) (supplementary information).

Differentially methylated regions (DMR)

The probe level approach was complemented using a regional analysis, which considers DNA methylation at clusters of neighbouring CpG sites throughout the epigenome. This approach is more statistically powerful and arguably more biologically plausible; neighbouring CpG sites are assumed to exert similar biological functions. We used the DMRff method (41) to identify differentially methylated regions (DMRs). Cohorts were supplied with an R script to conduct the DMR analysis using their own data (https://github.com/ammegandchips/Prenatal_Caffeine/blob/master/dmrff.mat.caff.EWAS.cohorts.r). Probes were annotated to the human reference genome version 19, build 37h using the annotation data available from the R-package *meffil* (42).

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Meta-analysis

Probe level meta-analysis

Results were meta-analysed with fixed effect estimates weighted by the inverse of the variance using the software METAL (43). Multiple testing was accounted for using a 5% false discovery rate (FDR) (44). Meta-analysed results were scrutinised in a similar manner as the individual cohort results and leave-one-cohort-out analysis using the R package *metafor* (45) was performed on the CpGs that showed evidence to be associated with maternal caffeine consumption. Results were deemed to be driven by a single cohort (and therefore to “fail” the leave-one-out test), if the meta-analysis effect estimate changed direction, moved towards the null by more than 20%, or had a confidence interval that included 0 after removal of a single cohort.

DMR meta-analysis

DMR cohort results were meta-analysed using an inverse-variance weighted fixed effects approach using the *dmrff.meta* function in the DMRFF R package (41). We defined a DMR as a region with at least two CpG sites with the same direction of effect and a Bonferroni adjusted P-value ($P_{\text{Bonferroni}} < 0.05$).

Causal inference and sensitivity analyses

Beverage-specific effects of the meta-analysed probe-level and DMR results were investigated by comparing the congruence between associations found using different sources of caffeine (that could have different confounding structures). Whereas high congruence between results of the different caffeine models (in terms of CpG site hits and/or genes annotated to CpG sites found in each model) would provide evidence for caffeine being the causal agent driving effects, beverage-specific effects would indicate that factors other than caffeine are driving associations.

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

To find out which gene pathways are linked to the CpG sites of the caffeine-associated DMRs, a gene ontology analysis was run using the R package *missMethyl* (46). We tested enrichment of gene ontology (GO) categories and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways.

Results

Sample characteristics

Maternal caffeine consumption during pregnancy

In all cohorts, most mothers (80-94%) consumed at least some caffeine during weeks 1 to 28 of gestation, with a weighted mean of 85 mg/day over all cohorts, but with large variation within and between cohorts (weighted average SD = 82 mg/day, Table 1).

Approximately 14% of mothers in the total sample consumed more caffeine than the commonly recommended caffeine limit of 200 mg/day. Across all cohorts, coffee and tea were the most common sources of caffeine, with coffee being the most common source in all cohorts except for the United Kingdom based cohorts ALSPAC and BiB, where the most common source was caffeinated tea (Table 1).

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Table 1. Overview of daily maternal caffeine consumption during pregnancy in the individual cohorts

Cohort (N)	Weeks of gestation of dietary assessment	Mean total daily caffeine intake (SD)	N users (%) > 200 mg of daily caffeine intake*	Mean daily intake of coffee (SD)	Mean daily intake of tea (SD)	Mean daily intake of cola (SD)
ALSPAC (n = 729)	18	135 (94)	197 (27%)	52 (69)	72 (58)	3 (6)
BIB (Asian; n = 353)	26-28	49 (47)	5 (2%)	11 (33)	41 (34)	12 (20)
BIB (white EU; n = 306)	26-28	112 (105)	50 (19%)	46 (65)	66 (63)	15 (22)
Generation R (n = 798)	18-25	115 (96)	132 (20%)	118 (78)	57 (62)	Not available
INMA (n = 378)	12	111 (129)	26 (8%)	79 (120)	25 (46)	7 (12)
EDEN (n = 156)	24-28	37 (43)	1 (<1%)	25 (42)	6 (15)	6 (9)
MoBa1 (N = 999)	22	105(106)	108 (11%)	70 (106)	18 (27)	14 (27)
Total and % or M and SD** (n = 3,724)	-	85 (82)	519 (14%)	46 (65)	26 (35)	5 (11)

*Note. M = Mean. SD = Standard deviation. * Mothers were grouped as users of caffeine if they indicated to consume more than zero cups of coffee, tea or cola. caffeine content in milligrams per day. ** In the Total row, average caffeine content was calculated by weighting by the inverse variance for each cohort. EU = European.*

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Demographics

A general overview over the demographics of the individual cohorts can be found in Table 2. Except for mothers from BiB, cohorts included slightly more mothers with higher (high school diploma or above) than lower educational attainment. Around 15% of mothers smoked after the 2nd trimester of pregnancy. Mothers who consumed caffeine during pregnancy were about twice as likely to have smoked during pregnancy (20%), compared to mothers who did not consume caffeine during pregnancy (11%; supplementary information, Table S1). Furthermore, mothers who consumed caffeine were more likely to already have children (49%) compared to mothers who did not consume caffeine during pregnancy (33%; supplementary information, Table S1).

Table 2. Overview of demographic information across the individual cohorts

Cohort	Country and ancestry	DNA methylation array	N high maternal SEP* (%)	M maternal age (SD)	N Maternal smoking (%)**	N parity > 0 (%)	M BMI (3.63)	M gestational age in weeks (1.52)
ALSPAC (N = 729)	United Kingdom; Northern European	450k	375 (51%)	29.79 (4.39)	77 (11%)	381 (52.3%)	22.79 (3.63)	39.53 (1.52)
BiB (Asian; N = 353)	United Kingdom; Pakistani	EPIC	146 (41%)	28.21 (5.37)	9 (3%)	249 (70.5%)	25.75 (5.23)	39.17 (1.52)
BiB (White EUR; N = 306)	United Kingdom; Northern European	EPIC	125 (41%)	26.98 (6.15)	93 (30%)	159 (52%)	27.10 (6.48)	39.29 (1.88)
Generation R (N = 798)	The Netherlands; Northern European	450k	458 (57%)	30.15 (4.95)	109 (14%)	95 (58.6%)	23.06 (3.64)	40.20 (1.48)
INMA (N = 378)	Spain; Southern European	450k	277 (73%)	31.55 (4.07)	53 (14%)	161 (42.6)	23.79 (4.44)	41.06 (1.34)
EDEN (N =162)	France; Southern & Northern European	450k	113 (70%)	31.94 (4.10)	26 (16%)	324 (40.6)	23.52 (4.64)	39.51 (1.33)
MoBa1 (N = 999)	Norway; Northern European	450k	761 (76%)	29.93 (4.35)	287 (29%)	580 (58.1)	24.02 (4.18)	39.95 (1.56)
Total or M (N = 3,725)	-	-	2,255 (61%)	30.01(4.60)	654 (18%)	1949 (52%)	23.61 (4.12)	39.94 (1.51)

Note. M = Mean. SD = Standard deviation. In the Total row, average caffeine content was calculated by weighting by the inverse variance for each cohort. * High maternal socio-economic position (SEP): maternal education >= high school diploma. ** continued smoking during pregnancy. Parity = one or more previous pregnancies. EU = European.

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Association between maternal caffeine consumption and offspring cord blood DNA methylation

Results of the quality control checks for the individual cohort results can be found in supplementary information, Figures S1 to S15. The overall EWAS meta-analysis results of the caffeine models can be found in Table 3. After adjusting for multiple testing, one CpG site (cg19370043, nearest gene *PRRX1*) was negatively associated with total maternal caffeine consumption (estimate = -2.18×10^{-05} ; 95% CI: -2.98×10^{-05} to -1.37×10^{-05} ; $P = 1.32 \times 10^{-07}$) and one CpG sites (cg14591243, nearest gene *STAG1*) with caffeine consumed from cola in mg/day (estimate = 2.78×10^{-05} , 95% CI: 2.779×10^{-05} to 2.781×10^{-05} , $P = 5.59 \times 10^{-09}$) (Table 3). Only the total maternal caffeine-associated CpG site survived the leave-one-out analysis. For the cola-associated CpG site, the leave-one-out analysis indicated that that MoBa was driving the effect (supplementary information, Figure S16 and S17).

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Table 3. A summary of results of each EWAS model from the probe-level analysis

Model*	CpGs with FDR-corrected P-value <0.05	CpGs surviving leave-one-out analysis	Meta-analysis sample size	Genomic inflation factor (λ)**
Any vs. no caffeine				
All offspring (minimally adjusted) *	0	n.a.	3731	0.97
All offspring (adjusted for covariates)	0	n.a.	3731	0.97
Female offspring (adjusted for covariates)	0	n.a.	1797	0.99
Male offspring (adjusted for covariates)	0	n.a.	1934	1.00
All offspring (adjusted for covariates and gestational age)	0		3731	0.97
Caffeine in mg/day				
All offspring (minimally adjusted)*	33	n.a.	3731	1.03
All offspring (adjusted for covariates)	1	1 (100%)	3731	1.00
Female offspring (adjusted for covariates)	0	n.a.	1797	1.00
Male offspring (adjusted for covariates)	0	n.a.	1934	1.04
All offspring (adjusted for covariates and gestational age)	0		3731	0.99
Caffeine from coffee				
All offspring (adjusted for covariates)	0	n.a.	2779	1.02
Caffeine from tea				
All offspring (adjusted for covariates)	0	n.a.	3477	1.00
Caffeine from cola				
All offspring (adjusted for covariates)	1	0	2610	1.00

*Note. * only adjusted for estimated cell counts and 20 surrogate variables. Covariates: maternal age, maternal smoking, maternal parity, maternal education, maternal BMI, estimated cell counts and 20 surrogate variables. ** The genomic inflation factor (λ) estimates the extent of bulk inflation of EWAS P-values and the excess false positive rate. 1 = no inflation; > 1 some evidence of inflation.*

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Beverage-specific effects

Figure 1 displays DNA methylation at the two CpG sites discovered in the probe-level analysis across the different caffeine models. If these associations were truly driven by caffeine exposure, we would expect to see similar associations of these CpG sites across the different beverage models. The association between DNA methylation at cg19370043 and maternal total caffeine consumption appears to be mostly driven by coffee and not by caffeine from tea or cola (Figure 1). For coffee intake specifically, the P-value was < 0.05 , but the association with coffee did not survive adjustment for multiple testing in the EWAS meta-analysis, probably because of the lower statistical power to detect small effects in the coffee compared to total caffeine analysis (n total caffeine = 3,731 vs. n coffee = 2,779; because the total caffeine model included mothers with missing beverage-specific data). The association between DNA methylation at cg14591243 appears to be specific to cola consumption rather than general caffeine consumption.

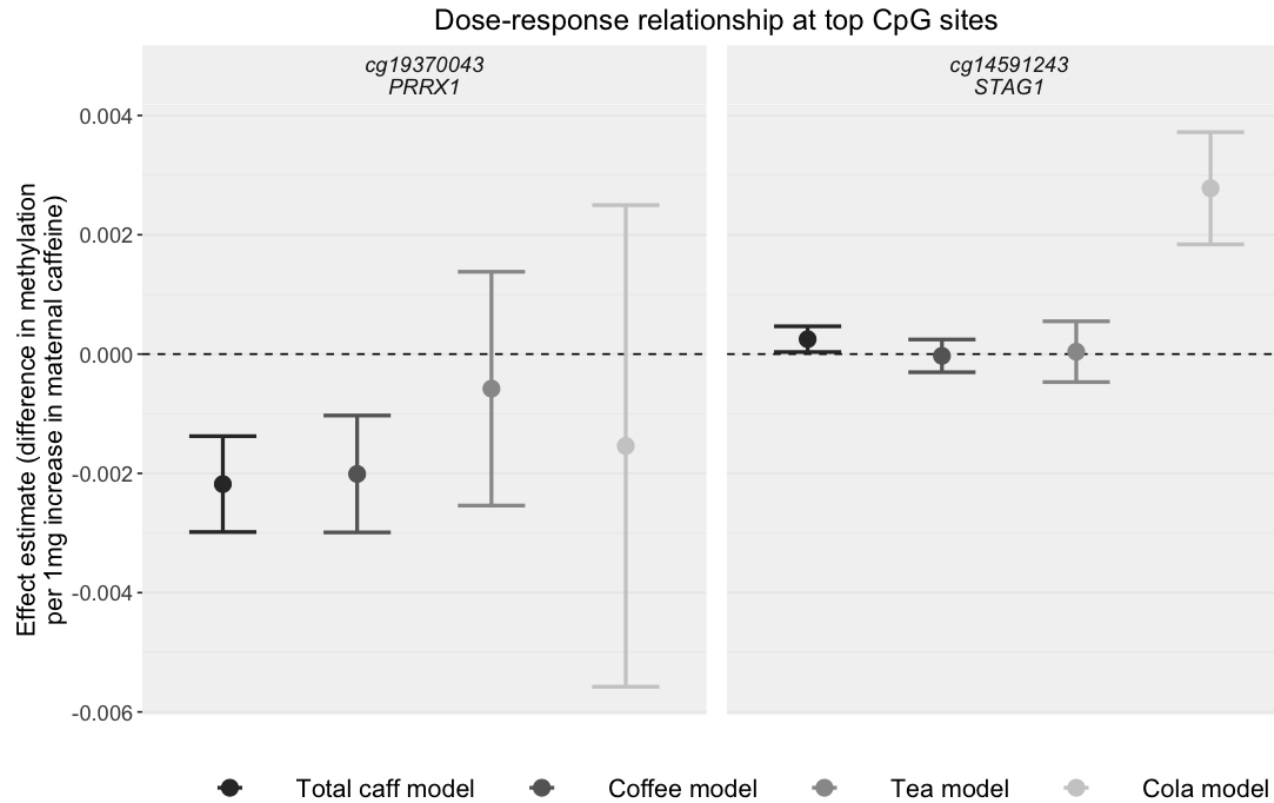


Figure 1. Effect size estimates at the top CpG sites found in the probe-level analysis. Total caff model = total caffeine. Coffee model = caffeine from coffee, Tea model = caffeine from tea, Cola model = caffeine from cola. Error bars represent 95% Confidence Intervals.

PRENATAL CAFFEINE AND OFFSPRING DNA METHYLATION

Differentially methylated regions (DMR) meta-analysis

The regional meta-analysis implemented using DMRff detected 22 differentially methylated regions for total maternal caffeine consumption with at least two and a maximum of 15 consecutive CpG sites ($P_{\text{Bonferroni}} < 0.05$; supplementary information, Table S2). The strongest evidence was found at a region on chromosome 17, with 7 consecutive CpG sites (chr17: 58499679-58499911; estimate = -3.77×10^{-05} ; SE = 5.02×10^{-06} ; $P_{\text{Bonferroni}} = 1.42 \times 10^{-10}$; supplementary information, Table S2). In the any vs. no maternal caffeine consumption model, there was evidence for 11 DMRs (supplementary information, Table S3). The strongest evidence was found at a region on chromosome 6, with 10 consecutive CpG sites (Chr6:31734147-31734554; estimate = -9.44^{-03} ; SE = 1.37^{-03} ; $P_{\text{Bonferroni}} = 1.93^{-06}$). DMR analyses from the individual sources of caffeine revealed 12 DMRs for caffeine consumed from coffee (supplementary information, Table S4), 18 DMRs for caffeine from tea (supplementary results, Table S5), and 14 DMRs for caffeine consumption from cola (supplementary information, Table S6) during pregnancy. The analyses from the sex-stratified models showed evidence for total maternal caffeine being associated with 12 DMRs in cord blood in female sex offspring (supplementary information, Table S7) and 18 DMRs in male sex offspring (supplementary information, Table S8).

For each pairwise combination of models, we calculated the percentage overlap of CpG sites (or closest gene) by dividing the overlap of CpG sites (or genes) between two models by the sum of the models' unique CpG sites/genes (e.g., percentage crossover any caffeine and total caffeine models: $7/(167 + 63 - 7) = 0.03 * 100 = 3\%$). There was very little overlap in CpG sites (range percentage overlap: 0-12%) or annotated genes (% overlapping genes: 0-11%) between the different models (supplementary information, Table S9).

Functional analysis of DMRs

Neither the functional categories defined by GO terms nor any of the KEGG pathways showed evidence for enrichment in genes annotated to CpG sites in the caffeine-associated DMRs (all FDR adjusted P -values > 0.05). The top 5 KEGG pathways and GO terms with the strongest evidence according to the smallest P -values for each list of CpGs in the caffeine-DMRs are available in supplementary information, Table S10.

Discussion

Summary & interpretation of findings

We investigated the association of self-reported maternal caffeine consumption during pregnancy with offspring cord blood DNA methylation using data from six international birth cohorts. For the EWAS meta-analysis, probe-level and regional DMR analyses were applied as hypothesis-free approaches to detect associations between maternal caffeine phenotypes and differential methylation levels in cord blood. We compared caffeine consumption during pregnancy across countries and caffeinated drinks, which reduced the potential for cultural confounding, and analysed the CpG sites of the maternal caffeine associated DMRs for their biological function. Results of these analyses show little converging evidence between the associations of the different sources of caffeine, indicating that the associations that we observed are most likely explained by other factors than caffeine exposure.

The probe-level analysis indicated that differences in DNA methylation at two CpG sites were associated with maternal caffeine intake (one with maternal total caffeine consumption and one with maternal caffeine consumption from cola); both showed small effect estimates. The coefficients from the regression analyses represent the change in offspring cord blood DNA methylation at a given CpG site per 1 mg/day increase in maternal caffeine consumption. Putting these results into real life context,

and assuming causality and linearity of effects, if the recommended limit of caffeine consumption during pregnancy were doubled from 200 mg/day to 400 mg/day, this would only be associated with 0.4% reduction in DNA methylation at cg19370043. This effect size is in line with the small effect sizes found in the EWAS meta-analysis of adult personal caffeine consumption on DNA methylation by Karabegović and colleagues (10), where an additional cup of coffee (=57mg) was associated with a 0.2% decrease in peripheral blood DNA methylation at a CpG site near *AHRR*, which would be equivalent to a 0.7% decrease in DNA methylation per 200 mg/day of caffeine (0.2% / 57 mg of caffeine per cup of coffee x 200). These estimated effect sizes appear to be much smaller than the estimated effect of smoking, which is the lifestyle exposure with the strongest effect on DNA methylation discovered to date. Sustained smoking during pregnancy was associated with changes of up to ~7% decrease in offspring cord-blood DNA methylation at the *AHRR* gene (36). As acknowledged by Karabegović et al, because their smoking adjustment did not include the amount of smoking or duration of smoking, the coffee-associated DNA methylation differences might be explained by residual confounding by smoking (10).

In the regional analyses, we identified 12-22 DMRs for each of the caffeine models. Yet, lack of congruence of associations across models, which was evident in the probe-level and regional analysis, provided evidence for beverage-specific effects instead of the effects being driven by caffeine (which is common to all included beverages).

Strengths & limitations

This was the first international EWAS meta-analysis investigating associations between offspring DNA methylation and caffeine from coffee, tea, and cola during pregnancy. A major strength of this study is the consideration of the effects of other common sources of caffeine besides coffee. Consumption of the different sources of caffeine might be

differentially socially patterned, allowing capturing a larger spectrum of the caffeine consuming population. For instance, British and non-European ethnicities consume more caffeinated tea than coffee (6,47). Further, it is indicated that the main source of caffeine might change during pregnancy, with even habitual coffee drinkers preferring caffeinated tea to coffee during pregnancy (48,49). Last, in contrast to previous research, this study assessed maternal caffeine consumption through mg/day instead of cups per day, which is a useful approach to isolate the effect of caffeine, allow comparison between different caffeinated beverages, and a more fine-tuned assessment of the effects of different caffeine dosages. We also adjusted for maternal smoking, an important potential confounder in analyses of caffeine intake, along with other potential confounding variables.

The findings should be considered in the light of the following limitations. Caffeine assessment in the meta-analysis relied on self-report, which might be underestimated (50) or underreported during pregnancy because of social stigma around maternal health behaviours (51,52). However, this would be more obvious for more recent cohorts and less and less likely for older cohorts, such as ALSPAC, due to awareness of the potential toxicity of caffeine in pregnancy emerging only recently. The questions used to assess caffeine consumption in the cohorts only allowed for rough estimations of maternal caffeine consumption (53), which could reduce power. Generalizability of the results to other populations might be limited by examining only first trimester consumption, the relatively low caffeine intake and the tendency for birth cohorts to enrol more advantaged families(16). Effects of caffeine exposure during pregnancy on offspring cord blood DNA methylation were only assessed at regions available on the 450k array, which only covers around 2% of CpG sites of the entire epigenome (54). Thus, differentially methylated CpGs or regions not covered by the array may have been missed in this study. We only assessed offspring DNA methylation in blood. There is

some evidence suggesting that DNA methylation levels in blood might be able to proxy for DNA methylation levels in other tissues, yet we cannot rule out that maternal caffeine during pregnancy might be influencing DNA methylation differentially in other tissue types (55). Finally, this study indicates that, if maternal caffeine consumption influences cord-blood DNA methylation, the effect is likely to be small. Although our meta-analysis maximised sample sizes, even larger sample sizes with more variable levels of caffeine consumption may be required to detect small effects of prenatal caffeine exposure on offspring DNA methylation.

Future research & conclusion

Future research should aim to use a more accurate assessment of caffeine consumption during pregnancy by considering differing types of coffee, brewing times, and cups sizes and/or assessing biomarkers of caffeine such as plasma concentrations of the caffeine metabolite paraxanthine (56). Triangulation strategies may be applied to disentangle confounded from causal effects of caffeine exposure during pregnancy on offspring DNA methylation. These might include: Further exploring the different confounding structures of various caffeinated beverages and considering individual differences in the maternal metabolism of caffeine. Maternal caffeine metabolism might influence intensity of exposure during pregnancy and might change the effect of caffeine on offspring DNA methylation. For instance, studies could conduct analyses using genetic variants that account for differences in caffeine metabolism (57) and/or consider pre-pregnancy caffeine consumption to account for differences in the tolerance to effects of caffeine during pregnancy (53). Also, more assessments of prenatal paternal caffeine consumption would enable the conduction of negative control analyses to investigate intrauterine effects (58,59), as well as investigating the effects of paternal caffeine consumption prior to pregnancy and its effect on offspring DNA methylation in its own right (59).

In conclusion, results of this large scale EWAS meta-analysis indicate little evidence for a strong effect of maternal caffeine consumption during the second trimester of pregnancy on offspring cord blood DNA methylation.

References

1. Reyes CM, Cornelis M. Caffeine in the diet: country-level consumption and guidelines. *Nutrients*. 2018;10(11):1772.
2. Grosso LM, Bracken MB. Caffeine Metabolism, Genetics, and Perinatal Outcomes: A Review of Exposure Assessment Considerations during Pregnancy. *Annals of Epidemiology*. 2005 Jun 1;15(6):460–6.
3. EFSA Panel on Dietetic Products, Nutrition, Allergies (NDA). Scientific Opinion on the safety of caffeine. *EFSA Journal*. 2015;13(5):4102.
4. Hammerton G, Munafò MR. Causal inference with observational data: the need for triangulation of evidence. *Psychol Med*. 2021 Mar 8;1–16.
5. Miyake Y, Tanaka K, Okubo H, Sasaki S, Arakawa M. Maternal caffeine intake in pregnancy is inversely related to childhood peer problems in Japan: The Kyushu Okinawa Maternal and Child Health Study. *Nutritional Neuroscience*. 2019 Nov 2;22(11):817–24.
6. Treur JL, Taylor AE, Ware JJ, McMahon G, Hottenga JJ, Baselmans BML, et al. Associations between smoking and caffeine consumption in two European cohorts: Smoking and caffeine consumption. *Addiction*. 2016 Jun;111(6):1059–68.
7. Sengpiel V, Elind E, Bacelis J, Nilsson S, Grove J, Myhre R, et al. Maternal caffeine intake during pregnancy is associated with birth weight but not with gestational length: results from a large prospective observational cohort study. *BMC Medicine*. 2013 Feb 19;11(1):42.
8. Felix JF, Joubert BR, Baccarelli AA, Sharp GC, Almqvist C, Annesi-Maesano I, et al. Cohort Profile: Pregnancy And Childhood Epigenetics (PACE) Consortium. *International Journal of Epidemiology*. 2018 Feb 1;47(1):22–23u.
9. Bird A. Perceptions of epigenetics. *Nature*. 2007 May 1;447(7143):396–8.
10. Karabegović I, Portilla-Fernandez E, Li Y, Ma J, Maas SCE, Sun D, et al. Epigenome-wide association meta-analysis of DNA methylation with coffee and tea consumption. *Nat Commun*. 2021 May 14;12(1):2830.
11. Buscariollo DL, Fang X, Greenwood V, Xue H, Rivkees SA, Wendler CC. Embryonic Caffeine Exposure Acts via A1 Adenosine Receptors to Alter Adult Cardiac Function and DNA Methylation in Mice. *PLoS One* [Internet]. 2014 Jan 27 [cited 2020 Feb 27];9(1). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3903656/>

12. Fang X, Mei W, Barbazuk WB, Rivkees SA, Wendler CC. Caffeine exposure alters cardiac gene expression in embryonic cardiomyocytes. *Am J Physiol Regul Integr Comp Physiol*. 2014 Dec 15;307(12):R1471–87.
13. Wu DM, He Z, Ma LP, Wang LL, Ping J, Wang H. Increased DNA methylation of scavenger receptor class B type I contributes to inhibitory effects of prenatal caffeine ingestion on cholesterol uptake and steroidogenesis in fetal adrenals. *Toxicol Appl Pharmacol*. 2015 Jun 1;285(2):89–97.
14. Xu D, Zhang B, Liang G, Ping J, Kou H, Li X, et al. Caffeine-Induced Activated Glucocorticoid Metabolism in the Hippocampus Causes Hypothalamic-Pituitary-Adrenal Axis Inhibition in Fetal Rats. *PLoS One* [Internet]. 2012 Sep 6 [cited 2020 Feb 27];7(9). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3435254/>
15. 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. *Int J Epidemiol*. 2013 Feb 1;42(1):111–27.
16. 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. *Int J Epidemiol*. 2013 Feb;42(1):97–110.
17. Wright J, Small N, Raynor P, Tuffnell D, Bhopal R, Cameron N, et al. Cohort Profile: The Born in Bradford multi-ethnic family cohort study. *International Journal of Epidemiology*. 2013 Aug 1;42(4):978–91.
18. Raynor P, Born in Bradford Collaborative Group. Born in Bradford, a cohort study of babies born in Bradford, and their parents: protocol for the recruitment phase. *BMC Public Health*. 2008 Sep 23;8:327.
19. Kooijman MN, Kruihof CJ, van Duijn CM, Duijts L, Franco OH, van IJendoorn MH, et al. The Generation R Study: design and cohort update 2017. *European Journal of Epidemiology*. 2016 Dec;31(12):1243–64.
20. Kruihof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CCW, et al. The Generation R Study: Biobank update 2015. *European Journal of Epidemiology*. 2014;29(12):911–27.
21. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, et al. Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol*. 2016 Apr;45(2):382–8.

22. Guxens M, Ballester F, Espada M, Fernández MF, Grimalt JO, Ibarluzea J, et al. Cohort Profile: the INMA--Infancia y Medio Ambiente--(Environment and Childhood) Project. *Int J Epidemiol*. 2012 Aug;41(4):930–40.
23. Heude B, Forhan A, Slama R, Douhaud L, Bedel S, Saurel-Cubizolles MJ, et al. Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *Int J Epidemiol*. 2016 Apr 1;45(2):353–63.
24. Thompson FE, Subar AF. Dietary Assessment Methodology. In: *Nutrition in the Prevention and Treatment of Disease* [Internet]. Elsevier; 2017 [cited 2020 Feb 24]. p. 5–48. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780128029282000011>
25. Farrow A, Shea KM, Little RE. Birthweight of term infants and maternal occupation in a prospective cohort of pregnant women. The ALSPAC Study Team. *Occupational and environmental medicine*. 1998;55(1):18–23.
26. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013 Feb 1;8(2):203–9.
27. Gervin K, Page CM, Aass HCD, Jansen MA, Fjeldstad HE, Andreassen BK, et al. Cell type specific DNA methylation in cord blood: A 450K-reference data set and cell count-based validation of estimated cell type composition. *Epigenetics*. 2016 Sep;11(9):690–8.
28. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* [Internet]. 2012 Dec [cited 2019 Apr 4];13(1). Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-86>
29. Yousefi P, Huen K, Davé V, Barcellos L, Eskenazi B, Holland N. Sex differences in DNA methylation assessed by 450 K BeadChip in newborns. *BMC Genomics* [Internet]. 2015 Nov 9 [cited 2020 Mar 3];16. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4640166/>
30. Merid SK, Novoloaca A, Sharp GC, Küpers LK, Kho AT, Roy R, et al. Epigenome-wide meta-analysis of blood DNA methylation in newborns and children identifies numerous loci related to gestational age. *Genome Medicine*. 2020 Mar 2;12(1):25.
31. York TP, Latendresse SJ, Jackson-Cook C, Lapato DM, Moyer S, Wolen AR, et al. Replicated

umbilical cord blood DNA methylation loci associated with gestational age at birth. *Epigenetics*. 2020 Nov 1;15(11):1243–58.

32. Bakker R, Steegers EA, Obradov A, Raat H, Hofman A, Jaddoe VW. Maternal caffeine intake from coffee and tea, fetal growth, and the risks of adverse birth outcomes: the Generation R Study. *Am J Clin Nutr*. 2010 Jun 1;91(6):1691–8.

33. Hoyt AT, Browne M, Richardson S, Romitti P, Druschel C. Maternal Caffeine Consumption and Small for Gestational Age Births: Results from a Population-Based Case–Control Study. *Matern Child Health J*. 2014 Aug;18(6):1540–51.

34. Elwert F, Winship C. Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annu Rev Sociol*. 2014 Jul;40:31–53.

35. Sharp GC, Alfano R, Ghantous A, Urquiza J, Rifas-Shiman SL, Page CM, et al. Paternal body mass index and offspring DNA methylation: findings from the PACE consortium. *Int J Epidemiol*. 2021 Jan 29;

36. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet*. 2016 Apr 7;98(4):680–96.

37. Tukey JW. *Exploratory data analysis*. Vol. 2. Reading, MA; 1977.

38. Leek JT, Johnson WE, Parker HS, Fertig EJ, Jaffe AE, Storey JD, et al. *sva: Surrogate Variable Analysis*. 2019.

39. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. *limma* powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*. 2015;43(7):e47.

40. Van der Most PJ, Küpers LK, Snieder H, Nolte I. QCEWAS: automated quality control of results of epigenome-wide association studies. *Bioinformatics*. 2017 Apr 15;33(8):1243–5.

41. Suderman M, Staley JR, French R, Arathimos R, Simpkin A, Tilling K. *dmrff: identifying differentially methylated regions efficiently with power and control*. *bioRxiv* [Internet]. 2018 Dec 31 [cited 2019 Feb 4]; Available from: <http://biorxiv.org/lookup/doi/10.1101/508556>

42. Suderman M, Hemani G, Min JL. *meffil: Efficient algorithms for DNA methylation* [Internet]. 2019. Available from: <https://github.com/perishky/meffil>

43. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association

scans. *Bioinformatics*. 2010 Sep 1;26(17):2190–1.

44. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995 Jan;57(1):289–300.

45. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*. 2010;36(3):1–48.

46. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analysing methylation data from Illuminas HumanMethylation450 platform. *Bioinformatics*. 2015;btv560.

47. The Coffee and Caffeine Genetics Consortium, International Parkinson’s Disease Genomics Consortium (IPDGC), North American Brain Expression Consortium (NABEC), UK Brain Expression Consortium (UKBEC), Cornelis M, Byrne EM, et al. Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Molecular Psychiatry*. 2015 May;20(5):647–56.

48. Chen L, Bell EM, Browne ML, Druschel CM, Romitti PA. Exploring Maternal Patterns of Dietary Caffeine Consumption Before Conception and During Pregnancy. *Matern Child Health J*. 2014 Dec;18(10):2446–55.

49. Lawson CC, LeMasters GK, Wilson KA. Changes in caffeine consumption as a signal of pregnancy. *Reproductive Toxicology*. 2004 Jul 1;18(5):625–33.

50. Schreiber GB, Maffeo CE, Robins M, Masters MN, Bond AP. Measurement of coffee and caffeine intake: Implications for epidemiologic research. *Preventive Medicine*. 1988 May 1;17(3):280–94.

51. Sharp GC, Lawlor DA, Richardson SS. It’s the mother!: How assumptions about the causal primacy of maternal effects influence research on the developmental origins of health and disease. *Social Science & Medicine*. 2018 Sep;213:20–7.

52. Murphy C, Brown T, Trickey H, Sanders J, Blaylock R, Dean C, et al. It remains unclear whether caffeine causes adverse pregnancy outcomes; but naive policy recommendations could cause harm [Letter to the editor] [Internet]. 2020. Available from: <https://ebm.bmj.com/content/early/2020/09/01/bmjebm-2020-111432.responses#it-remains-unclear-whether-caffeine-causes-adverse-pregnancy-outcomes-but-naive-policy-recommendations-could-cause-harm>

53. van Dam RM, Hu FB, Willett WC. Coffee, Caffeine, and Health. *Campion EW, editor. New England*

Journal of Medicine. 2020 Jul 23;383(4):369–78.

54. Lövkvist C, Dodd IB, Sneppen K, Haerter JO. DNA methylation in human epigenomes depends on local topology of CpG sites. *Nucleic Acids Res.* 2016 Jun 20;44(11):5123–32.

55. Walton E, Relton CL, Caramaschi D. Using Openly Accessible Resources to Strengthen Causal Inference in Epigenetic Epidemiology of Neurodevelopment and Mental Health. *Genes.* 2019 Mar;10(3):193.

56. Boylan SM, Cade JE, Kirk SFL, Greenwood DC, White KLM, Shires S, et al. Assessing caffeine exposure in pregnant women. *Br J Nutr.* 2008 Oct;100(4):875–82.

57. Cornelis M, Kacprowski T, Menni C, Gustafsson S, Pivin E, Adamski J, et al. Genome-wide association study of caffeine metabolites provides new insights to caffeine metabolism and dietary caffeine-consumption behavior. *Human Molecular Genetics.* 2016 Oct 3;ddw334.

58. Davey-Smith G. Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? *Basic & clinical pharmacology & toxicology.* 2008;102(2):245–56.

59. Easey KE, Sharp GC. The impact of paternal alcohol, tobacco, caffeine use and physical activity on offspring mental health: a systematic review and meta-analysis. *Reprod Health.* 2021 Oct 26;18:214.

Dose-response relationship at top CpG sites

