This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Schellhas, Laura

Title:
Smoking and Caffeine Consumption during Pregnancy and Offspring Mental Health
Disentangling Causal from Confounded Associations using Genetic and Epigenetic Methods

General rights
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

Smoking and Caffeine Consumption during Pregnancy and Offspring Mental Health: Disentangling Causal from Confounded Associations using Genetic and Epigenetic Methods

Laura Schellhas

July 2021

School of Psychological Science

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Life Sciences

Word count: 67,152

// Abstract

It is unclear whether observational associations between smoking and caffeine consumption during pregnancy and offspring mental health problems are causal or confounded. More robust evidence is needed to enable parents to make an informed choice about behaviours during pregnancy. Specifically, for caffeine consumption during pregnancy – a behaviour that is strongly correlated with smoking – pregnancy guidelines are unclear, leading to confusion and uncertainty for expecting parents. Understanding whether smoking and caffeine consumption during pregnancy are causally related to offspring mental health would strengthen the causal evidence base.

The identification of a molecular pathway that could explain how prenatal exposures become biologically embedded to increase the risk for offspring health outcomes could add to the evidence for a causal relationship. Despite a hypothesised role for DNA methylation in mediating the link between prenatal exposures and mental health outcomes, very few studies have tested this hypothesis. In this thesis, I applied genetically informed analyses, as well as examined DNA methylation as a potential biological mechanism, to understand whether prenatal smoking and caffeine exposure are causally related to mental health outcomes.

The results of this thesis indicate that observational associations between prenatal smoking exposure and externalising problems are likely to be confounded by a shared genetic liability between (maternal) smoking and (offspring) risk-seeking personality traits, in turn affecting offspring mental health outcomes. I found little evidence for associations between prenatal caffeine exposure and mental health outcomes. In line with these results, I found no evidence for a causal contribution of prenatal smoking- and caffeine-associated DNA methylation to offspring internalising problems. These results are in accordance with recent research, which indicates that, given the currently available technologies, DNA methylation may be more valuable as a biomarker for prenatal exposures and mental health outcomes than for understanding pathways to mental health problems.

Acknowledgements

I would like to thank my three supervisors Dr. Gemma Sharp, Dr. Luisa Zuccolo and Prof. Dr. Marcus Munafò for all their support and encouragement throughout the past years. To Gemma and Luisa, thank you for being such inspiring, female-scientist role models and for your positive feedback, that kept me motivated. To Gemma, thank you for always having my back and my best interest in mind, teaching me about epigenetics, taking your time during our weekly check-ins, your career advice, and emotional support. Thank you, Luisa, for your patience in teaching, encouragement to keep asking questions, and your pastoral support during the pandemic. Thank you, Marcus, for making me feel welcome in TARG, your direct and efficient communication – which was very helpful for me as a German moving to England for the first time – your responsiveness and advice, and for always staying on top of things. Thank you all, for helping me to set foot in the fields of genetics, epigenetics, and epidemiology; I could not have asked for a better team of supervisors!

Thank you TARG, for showing me how collaborative, encouraging, supportive, fun, and non-competitive research teams can be. Thank you DreamStream, especially Hannah and Robyn, for all your support in understanding genetic epidemiology. Thank you to the psychology PhD community, especially Maddy and Jaz, for always getting the group together, and your help in managing the PhD student life. Also, thank you to my office mates Daria, Mubs, and Katie – the pandemic made me appreciate our office chats, lunches, and walks even more.

I am grateful to the EU commission for funding my PhD and enabling me to travel around the world to connect with international scientists. I am very glad that I could do my PhD within the CAPICE project, which allowed me to work and become friends with so many brilliant, funny, interesting, and diverse young career researchers. I am especially grateful to have had Elis as another CAPICE student at my side from day one of this journey. Together, we built us homes in Bristol, Rotterdam, and Norway and got to travel to so many remarkable places – Without you it would have been only half the fun.

A special thank you for pandemic-related support goes to the members of the TARG virtual writing retreats and my virtual office mates Minita, Ole, and Nadja. Thank you for helping me to get words down on paper during isolation and lockdown.

Finally, I am very grateful to my supportive and loving family. Thank you Mama and Papa, for always believing in me, teaching me to not give up when things get challenging, and encouraging me to explore new places around the world, while letting me know that I can always come home anytime I want or need to. Thank you to my sisters Eva and Sophia, for helping to build my strong mind-set and always being there for me. Last but not least, a huge thank you to my partner Lasse, who has been my biggest supporter during this time. Thank you for your open-mindedness to move to Bristol with me, for all the pick-me ups and pep talks, for listening to my presentations, proof-reading my work, and providing me with so much emotional support throughout this time – I truly could not have done it without you!

Author’s Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: DATE:

Publications

Schellhas, L., Haan, E., Easey, K. E., Wootton, R. E., Sallis, H. M., Sharp, G. C., Munafò, M. R., & Zuccolo, L. (2021). Maternal and child genetic liability for smoking and caffeine consumption and child mental health: An intergenerational genetic risk score analysis in the ALSPAC cohort. *Addiction*. https://doi.org/10.1111/add.15521

Easey, K. E., Wootton, R. E., Sallis, H. M., Haan, E., **Schellhas, L**., Munafò, M. R., Timpson, N. J., & Zuccolo, L. (2021). Characterization of alcohol polygenic risk scores in the context of mental health outcomes: Within-individual and intergenerational analyses in the Avon Longitudinal Study of Parents and Children. *Drug and Alcohol Dependence*, *221*, 108654. https://doi.org/10.1016/j.drugalcdep.2021.108654

Sharp, G. C., **Schellhas, L.**, Richardson, S. S., & Lawlor, D. A. (2019). Time to cut the cord: Recognizing and addressing the imbalance of DOHaD research towards the study of maternal pregnancy exposures. *Journal of Developmental Origins of Health and Disease*, *10*(5), 509–512. https://doi.org/10.1017/S2040174419000072

Rajula, H. S. R., Manchia, M., Agarwal, K., Akingbuwa, W. A., Allegrini, A. G., Diemer, E., Doering, S., Haan, E., Jami, E. S., Karhunen, V., Leone, M., **Schellhas, L**., Thompson, A., van den Berg, S. M., Bergen, S. E., Kuja-Halkola, R., Hammerschlag, A. R., Järvelin, M. R., Leval, A., … Middeldorp, C. M. (2021). Overview of CAPICE-Childhood and Adolescence Psychopathology: Unravelling the complex etiology by a large Interdisciplinary Collaboration in Europe-an EU Marie Skłodowska-Curie International Training Network. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-020-01713-2

Murphy, C., Brown, T., Trickey, H., Sanders, J., Blaylock, R., Dean, C., Hennsessy, M., **Schellhas, L**., Sharp, G. C., Zuccolo, L., Munafò, M. R., Cairns, K., Booker, M., Tennant, P., Fisher, J., Lee, E., Williams, E., Duff, E., Petersen, I., & Marshall, E. (2020, September 3). *It remains unclear whether caffeine causes adverse pregnancy outcomes; but naive policy recommendations could cause harm [Letter to the editor].* 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

Plans for publishing and disseminating the research findings of this thesis:

Schellhas, L., Monasso, G., Felix, J., Fernández-Barrés, S., Pesce, G., Annesi-Maesano, I., Page, C., London, S., Zuccolo, L., Munafò, M.R., Sharp, G.C. (2021). Maternal caffeine consumption during pregnancy and offspring cord blood DNA methylation: a meta-analysis of six epigenome-wide association studies - Manuscript in preparation.

Schellhas, L., Lou, M., Cecil, C., Page, C, Ystrom, E., Havdahl, A., Zuccolo, L., Munafò, M.R., Sharp, G.C. (2021). Inspecting critical windows for the development of internalising problems during childhood: A meta-analysis of three epigenome-wide-association-studies - Manuscript in preparation.

Table of Contents

Table of Contents .... **i**

List of Tables .... **viii**

List of Figures .... **x**

List of Appendices .... **xiii**

List of Abbreviations .... **xix**

Chapter 1 – Introduction **1**

1.1 Chapter overview *1*

1.2 Developmental Origins of Health and Disease hypothesis and offspring mental health outcomes *1*

1.3 Mental health problems in childhood and adolescence *2*

1.4 Prevalence of offspring externalising and internalising problems *3*

1.5 Prevalence of smoking and caffeine consumption during pregnancy *5*

1.5.1 Smoking 5

1.5.2 Caffeine 7

1.6 Observational evidence of prenatal smoking and caffeine exposure and mental health outcomes *8*

1.6.1 Smoking 8

1.6.2 Caffeine 10

1.7 Biological effects of intrauterine exposure to smoking and caffeine on offspring mental health development *13*

1.7.1 Biological pathways of smoking and caffeine during pregnancy on offspring's brain development 13

1.7.2 Pregnancy related changes in tobacco and caffeine metabolism 16

1.7.3 Epigenetics as a pathway linking intrauterine exposure to smoking and caffeine on offspring mental health outcomes 17

1.8 Problems with observational evidence *19*

1.8.1 Selection bias 20

1.8.2 Measurement bias 21

1.8.3 Confounding bias 22

1.9 Why we need better evidence 27

1.9.1 Pregnancy guidelines 27

1.9.2 Mother blaming 29

1.9.3 Testing, rather than reinforcing, the DOHaD hypothesis 31

1.10 Causal Inference Techniques 35

1.10.1 Negative control studies 35

1.10.2 Mendelian Randomisation (MR) 37

1.10.3 Other genetically sensitive designs to infer causality 39

1.10.4 The meet-in-the-middle approach to establish causality 40

1.10.5 Triangulation 41

1.11 Chapter summary and conclusion 42

Chapter 2 – Methods and Samples 44

2.1 Chapter overview 44

2.2 Methods 44

2.2.1 Genetically informed analyses 44

2.2.2 Pleiotropy 50

2.3 Epigenome-wide association studies (EWAS) 52

2.3.1 Assessment of DNA methylation using Illumina microarrays 53

2.3.2 Biological variation in DNA methylation 55

2.3.3 Technical variation in DNA methylation assessment 61

2.3.4 Statistical analysis 62

2.3.5 Meta-analysis quality control checks (Step 5 and 6 of meta-EWAS pipeline) 67

2.3.6 Differentially methylated regions 68

2.4 Mendelian Randomziation applied to epigenetic data 68

2.4.2 One-sample and two-sample MR 70

2.5 Multiple-testing correction 72

2.6 Cohorts 73

2.6.1 Avon Longitudinal Study of Parents and Children (ALSPAC) 73

2.6.2 Generation Rotterdam (Generation R) 78

2.6.3 Norwegian Mother and Child Cohort Study (MoBa) 80

2.6.4 Born in Bradford (BiB) 82

2.6.5 INfancia y Medio Ambiente (Environment and Childhood; INMA) 84

2.6.6 Etude des Déterminants pré et post natals du développement et de la santé de l′ Enfant (EDEN) 85

2.7 Databases 87

2.7.1 Genecards 87

2.7.2 NHGRI-EBI GWAS Catalog 87

2.7.3 EWAS catalog 88

2.8 Chapter summary 89

Chapter 3 – Maternal and child genetic liability for smoking and caffeine consumption and child mental health: An intergenerational polygenic risk score analysis in the ALSPAC cohort 90

3.1 Chapter overview 92

3.2 Introduction 92

3.2.1 Intergenerational Mendelian Randomisation 94

3.2.2 Phenome-wide association studies (PheWAS) 98

3.2.3 Chapter Aims 99

3.3 Methods 101

3.3.1 Study population 101

3.3.2 Design 101

3.3.3 Phenotype data 102

3.3.4 Polygenic risk scores (PRS) 104

3.3.5 Statistical analysis 105

3.3.6 Multiple testing 106

3.4 Results 107

3.4.1 Maternal smoking and caffeine consumption 107

3.4.2 Aim 1: Validation of PRS during pregnancy 108

3.4.3 Aim 2: Comparison of intergenerational and childhood smoking initiation PRS analyses 111

3.4.4 Aim 2: Comparison of intergenerational and childhood caffeine PRS analyses 114

3.4.5 Supplementary analyses: Associations between maternal and offspring PRS and own mental health during and outside of pregnancy (mothers) and adolescence (offspring) 116

3.5 Discussion 118

3.5.1 Summary and interpretation of findings 118

3.5.2 Smoking PRS and mental health 118

3.5.3 Caffeine PRS and Mental Health 119

3.5.4 Smoking and caffeine consumption 121

3.5.5 Strengths and Limitations 122

3.5.6 Future research 123

3.5.7 Conclusion 124

Chapter 4 – Maternal caffeine consumption during pregnancy and offspring cord blood DNA methylation: a meta-analysis of six epigenome-wide association studies 125

4.1 Chapter overview 125

4.2 Introduction 126

4.3 Methods ... *127*

4.3.1 Meta-analysis of epigenome-wide association studies (EWAS) ... 127

4.3.2 Statistical analyses ... 129

4.4 Results ... *135*

4.4.1 Sample characteristics ... 135

4.4.2 Probe level meta-analysis ... 140

4.4.3 Differentially methylated regions (DMR) meta-analysis ... 151

4.4.4 Analysis of cell proportions ... 161

4.5 Discussion ... *162*

4.5.1 Summary & interpretation of findings ... 162

4.5.2 Strengths ... 166

4.5.3 Limitations ... 167

4.5.4 Future research ... 168

4.5.5 Conclusion ... 169

Chapter 5 – The development of childhood internalising problems: A meta-analysis of epigenome-wide-association-studies ... 170

5.1 Chapter overview ... *170*

5.2 Introduction ... *170*

5.2.1 Epigenetic studies of offspring anxiety and depressive symptoms ... 172

5.2.2 Epigenome-wide association studies of anxiety and depression ... 174

5.2.3 Summary ... 175

5.3 Methods ... *176*

5.3.1 Meta-analysis of epigenome-wide association studies ... 176

5.3.2 Statistical analyses ... 182

5.3.3 Candidate-gene wide analysis ... 183

5.4 Results ... *184*

5.4.1 Sample characteristics ... 184

5.4.2 Probe-level meta-analysis ... 190

5.4.3 Differentially methylated regions meta-analysis ... 193

5.4.4 Candidate gene look-up ... 199

5.4.5 Meta-analysis of cell proportions ... 202

5.5 Discussion ... 202

5.5.1 Summary and interpretation of findings ... 202

5.5.2 Biological pathways ... 204

5.5.3 Strengths ... 205

5.5.4 Limitations ... 206

5.5.5 Future studies ... 208

5.6 Conclusion ... 209

Chapter 6 – DNA methylation as a mediator of the association between smoking and caffeine consumption during pregnancy and offspring internalising problems ... 210

6.1 Chapter overview ... 210

6.2 Introduction ... 210

6.3 Methods ... 218

6.3.1 Data selection ... 218

6.3.2 Statistical analysis ... 221

6.4 Results ... 223

6.4.1 Enrichment of smoking and caffeine associated CpG sites in the internalising EWAS results ... 223

6.4.2 Clumping of the exposure instruments (cis-SNPs) ... 225

6.4.3 Two-sample MR of prenatal smoking-associated DNA methylation and childhood internalising problems, and adult depression and anxiety ... 226

6.4.4 Two-sample MR of prenatal caffeine-associated DNA methylation and childhood internalising problems, and depression and anxiety in adulthood 228

6.4.5 Two-sample MR of internalising problems-associated DNA methylation and childhood internalising problems, and depression and anxiety in adulthood 230

6.5 Discussion 230

6.5.1 Summary 230

6.5.2 Strengths 232

6.5.3 Limitations 232

6.5.4 Future directions for epigenetic research in mental health 233

6.6 Conclusion 234

Chapter 7 – Discussion 235

7.1 Chapter overview 235

7.2 Summary of thesis findings 235

7.3 Contribution of this thesis towards the existing DOHaD literature 241

7.3.1 Methodological strengths and limitations 241

7.4 The future of genetics and epigenetics for prevention and treatment of mental health problems 247

7.4.1 Polygenic risk scores 250

7.4.2 DNA methylation risk scores 251

7.5 Ethical considerations of the development of an epigenetic biosocial archive 256

7.6 Main conclusion 258

References 261

Appendices 326

List of Tables

Table 3.1 Availability of phenotypes included in the intergenerational and childhood analyses across the two generations 103

Table 3.2 Associations between smoking initiation polygenic risk scores and smoking phenotypes in mothers and offspring in adolescence. 109

Table 3.3 Associations between caffeine polygenic risk scores and daily caffeine intake in mothers and offspring 110

Table 4.1 Overview of maternal caffeine consumption during pregnancy in the individual cohorts 136

Table 4.2 Overview of maternal demographic information across the individual cohorts 138

Table 4.3 Demographic variables of the cohorts stratified by caffeine user status 139

Table 4.4 A summary of results of each EWAS model from the probe-level analysis .. 142

Table 4.5 CpG sites associated with maternal total caffeine consumption and caffeine consumption from cola with FDR-adjusted P-value < 0.05 143

Table 4.6 Looking-up CpG sites found to be associated with maternal caffeine consumption in the results of associations between paternal caffeine consumption and offspring DNA methylation. 145

Table 4.7 Summary of results of paternal caffeine models in ALSPAC 149

Table 4.8 Results of the maternal caffeine PRS analysis in ALSPAC 151

Table 4.9 Crossover of CpG sites and genes of the DMRs of the different caffeine models 158

Table 5.1 Maternal characteristics 184

Table 5.2 Offspring characteristics 187

Table 5.3 Correlation table of internalising problems and covariates 189

Table 5.4 Correlation between maternal smoking during pregnancy and offspring internalising problems 189

Table 5.5 A summary of results of each EWAS model from the probe-level analysis .. 192

Table 5.6 CpG sites associated with offspring DNA methylation and internalising problems at the age of 7 with FDR-adjusted P-value < 0.05 193

Table 5.7 Results from the meta-analysis of differentially methylated regions for the cord blood meta-analysis: Internalising problems age 3..194

Table 5.8 Results from the meta-analysis of differentially methylated regions for the cord blood meta-analysis: Internalising problems age 7..196

Table 5.9 Results from the meta-analysis of differentially methylated regions for the cross-sectional analysis: Internalising problems age 7 198

Table 6.1 Results of the two-sample MR of prenatal smoking-associated DNA methylation changes at birth and childhood internalising problems, and adult anxiety and depression...227

Table 6.2 Results of the two-sample MR analysis of prenatal caffeine-associated DNA methylation at birth on childhood internalising problems, and adult anxiety and depression 229

Table 7.1 Qualitative Summary of results from each study of this thesis 239

Table 7.2 Summary of strengths and limitations of each of the methods used in this thesis 245

List of Figures

Figure 1.1 Prevalence rates of mental health problems in young people from 1999 to 2017 (credit: Sadler et al., 2018). 4

Figure 1.2. Forest plot of the studies included in the meta-analysis by Duko et al. investigating prenatal smoking and the risk for offspring mood disorders (credit: Duko et al., 2020). 9

Figure 1.3. Risk for offspring mental health problems at the age of 11 in association with total maternal caffeine consumption at 15- and 30-weeks gestation (credit: Hvolgaard Mikkelsen et al., 2017). 11

Figure 1.4. Sunburst of studies published in the Journal of the Developmental Origins of Health and Disease and the proportions of their focus on maternal and paternal exposures (credit: Sharp et al., 2019). 33

Figure 2.1 Illustration of vertical and horizontal pleiotropy for assessing the effect of maternal smoking on offspring mental health outcomes 52

Figure 2.2 Illustration of DNA methylation patterns near a gene region (credit: Jones et al., 2018) 55

Figure 2.3 First six steps of the EWAS meta-analysis pipeline introduced by Sharp and colleagues (credit: Sharp et al., 2021). 63

Figure 2.4 Precision plots for cohort results 66

Figure 2.5 Illustration of a two-sample MR analysis investigating the causal effect of exposure-associated DNA methylation on a phenotype 71

Figure 3.1 Application of intergenerational MR to investigate the intrauterine effect of smoking on mental health outcomes (credit: adapted from Lawlor et al., 2017)..... 96

Figure 3.2 Illustration of offspring smoking PRS as a collider caused by maternal and paternal smoking PRS 97

Figure 3.3 Design overview of the comparison of the intergenerational analysis and the childhood analysis 101

Figure 3.4 Overview of the intake of the different sources of caffeine outside and during pregnancy 108

Figure 3.5 Comparison of phenotype associations with the smoking initiation polygenic risk scores in the intergenerational and childhood analysis 113

Figure 3.6 Comparison of phenotype associations with the caffeine polygenic risk scores in the intergenerational and childhood analysis 115

Figure 4.1 Visualisation of average maternal caffeine consumption during pregnancy across caffeine sources and cohorts 135

Figure 4.2 Effect size estimates at the top CpG sites found in the probe-level analysis. 144

Figure 4.3 Comparison of maternal and paternal caffeine EWAS effect estimates at the top CpG sites of the meta-analysis results in ALSPAC 147

Figure 4.4 Dot plots comparing offspring DNA methylation at the top 5 CpG sites of the any paternal caffeine models in ALSPAC 150

Figure 4.5 Candidate DMRs for total maternal caffeine consumption. 152

Figure 4.6 Candidate DMRs for any maternal caffeine consumption 153

Figure 4.7 Candidate DMRs for maternal coffee consumption 154

Figure 4.8 Candidate DMRs for maternal tea consumption 155

Figure 4.9 Candidate DMRs for maternal cola consumption 156

Figure 4.10 Candidate DMRs for female sex stratified models of total maternal caffeine consumption 160

Figure 4.11 Candidate DMRs for male sex stratified models of total maternal caffeine consumption 161

Figure 5.1 Flow chart of the sample sizes of the cord blood analyses. 178

Figure 5.2 Flow chart of the sample size of the cross-sectional analysis 178

Figure 5.3 Manhattan plot of the candidate gene look-up of the internalising problems age 3 ~ cord-blood results 199

Figure 5.4 Manhattan plot of the candidate gene look-up of the internalising problems age 7 ~ cord-blood results. significance level 200

Figure 5.5 Manhattan plot of the candidate gene look-up of the internalising problems age 7 ~ childhood peripheral blood results 200

Figure 6.1 Two-step MR to investigate offspring DNA methylation as a mediator of prenatal caffeine exposure on offspring internalising problems (credit: adapted from Relton & Davey Smith, 2012) 212

Figure 6.2 Illustration of offspring PRS as a collider of maternal and paternal caffeine PRS 213

Figure 6.3 Illustration of a two-sample MR analysis to investigate whether DNA methylation mediates the relationship of prenatal caffeine exposure on offspring internalising problems 215

Figure 6.4 QQ-Plots of the association of the EWAS meta-analysis of internalising problems at the top 5,000 prenatal smoking-associated CpG sites 224

Figure 6.5 QQ-Plots of the association of the EWAS meta-analysis of internalising problems at the top 5,000 prenatal caffeine-associated CpG sites. 225

Figure 7.1 Overview of analyses used in this thesis 238

Figure 7.2 Ancestry populations included in published GWAS (credit: Mills & Rahal, 2019) 251

Figure 7.3 Number of epigenetic epidemiology publications per year 253

List of Appendices

Appendix A **...326**

A1 List of phenotypes included in the study. 326

Appendix B **...331**

B1 Two-sample t-tests (two-sided) to test differences between participants with complete and partially missing genotype data...............................331

B2 Chi-Square test to test differences between participants with complete and partially missing genotype data...............................332

Appendix C **...333**

C1 Associations between the lifetime smoking PRS and smoking phenotypes in mothers during and outside of pregnancy and adolescents..............................333

Appendix D **...334**

D1 Associations between the maternal and offspring smoking initiation PRS and offspring phenotypes <10 years...............................334

Appendix E **...335**

E1 Associations between maternal and offspring lifetime smoking PRS and offspring phenotypes <10 years. 335

Appendix F **...336**

F1 Associations between maternal and offspring caffeine PRS and offspring phenotypes <10 years. 336

Appendix G **...337**

G1 Associations between the maternal and offspring smoking initiation PRS and phenotypes in mothers during and outside of pregnancy and adolescence............337

Appendix H **...341**

H1 Associations between maternal and offspring lifetime smoking PRS and phenotypes in mothers during and outside pregnancy and adolescence...............................341

Appendix I **...345**

I1 Associations between caffeine PRS and phenotypes in mothers during and outside of pregnancy and adolescence..............................345

Appendix J **349**

J1 Correlation between smoking, caffeine and alcohol PRS 349

Appendix K **350**

K1 ALSPAC definition of variables and covariates 350

Appendix L **356**

L1 Correlation plot ALSPAC 356

L2 Correlation plot BiB (Asian ethnicity) 356

L3 Correlation plot BiB (White European ethnicity) 357

L4 Correlation plot Generation R 357

L5 Correlation plot EDEN 358

L6 Correlation plot INMA 358

L7 Correlation plot MoBa1 359

Appendix M **360**

M1 ALSPAC QQ-Plots of caffeine models 360

M2 BiB (Asian ethnicity) QQ-Plots of caffeine models 361

M3 BiB (White European ethnicity) QQ-Plots of caffeine models 362

M4 Generation R QQ-Plots of caffeine models 363

M5 MoBa1 QQ-Plots of caffeine models 364

M6 EDEN QQ-Plots of caffeine models 365

M7 INMA QQ-Plots of caffeine models 366

Appendix N **367**

N1 Precision plots 367

Appendix O **368**

O1 Leave-one-out plot of the prenatal caffeine-associated CpG site (Cg19370043) .. 368

O2 Leave-one-out plot of the prenatal cola-associated CpG site (Cg12788467) 368

O3 Leave-one-out plot of the prenatal cola-associated CpG site (Cg14591243) 369

Appendix P **...370**

P1 Correlation matrix of the meta-analysed caffeine models 370

Appendix Q 371

Q1 Top five associations between maternal caffeine polygenic risk score (PRS) and offspring DNA methylation 371

Appendix R...372

R1 Results from the meta-analysis of differentially methylated regions for maternal total caffeine consumption 372

R2 Results from the meta-analysis of differentially methylated regions for any vs. no maternal caffeine consumption. 373

R3 Table H3 Results from the meta-analysis of differentially methylated regions for maternal coffee consumption. 374

R4 Results from the meta-analysis of differentially methylated regions for maternal tea consumption. 375

R5 Results from the meta-analysis of differentially methylated regions for maternal cola consumption. 376

R6 Results from the meta-analysis of differentially methylated regions for maternal caffeine consumption stratified by female sex...377

R7 Results from the meta-analysis of differentially methylated regions for maternal caffeine consumption stratified by male sex...378

Appendix S 379

S1 Top 5 GO terms and KEGG pathways for CpGs in DMRs (BP = biological process; MF = molecular function; CC = cell compartment) 379

Appendix T...382

T1 ALSPAC definition of variables and covariates..382

T2 Generation R definition of variables and covariates..383

T3 MoBa definition of variables and covariates 384

Appendix U **385**

U1 Correlation plot cord blood meta-analysis age 3 – ALSPAC 385

U2 Correlation plot cord blood meta-analysis age 3 – Generation 385

U3 Correlation plot cord blood meta-analysis age 3 – MoBa1 386

U4 Correlation plot cord blood meta-analysis age 3 – MoBa2 386

U5 Correlation plot cord blood meta-analysis age 7 – ALSPAC 387

U6 Correlation plot cord blood meta-analysis age 7 – Generation R 387

U7 Correlation plot cross-sectional meta-analysis age 7 – ALSPAC 388

U8 Correlation plot cross-sectional meta-analysis age 7 – Generation R 388

Appendix V **389**

V1 QQ-plots cord blood meta-analysis age 3 – ALSPAC 389

V2 QQ-plots cord blood meta-analysis age 3 – Generation R 390

V3 QQ-plots cord blood meta-analysis age 3 – MoBa1 391

V4 QQ-plots cord blood meta-analysis age 3 MoBa2 392

V5 QQ-plots of the cord blood meta-analysis age 7 – ALSPAC 393

V6 QQ-plots of the cord blood meta-analysis age 7 – Generation R 394

V7 QQ-plots of the childhood peripheral blood meta-analysis age 7 – ALSPAC 395

V8 QQ-plots of the childhood peripheral blood meta-analysis age 7 – Generation R 396

Appendix W **397**

W1 Precision plot of the cord blood meta-analysis age 3 397

W2 Precision plot of the cord blood meta-analysis age 7 398

W3 Precision plot of the childhood peripheral blood meta-analysis age 7 399

Appendix X **400**

X1 Correlation matrix of the cord blood probe-level meta-analysis results: Internalising problems age 3 400

X2 Correlation matrix of the cord blood probe-level meta-analysis results: Internalising problems age 7 401

X3 Correlation matrix of the childhood peripheral blood probe-level meta-analysis results: Internalising problems age 7 402

Appendix Y 403

Y1 QQ-plot of the cord blood probe-level meta-analysis results: Internalising problems age 3 403

Y2 QQ-plot of the cord blood probe-level meta-analysis results: Internalising problems age 7 404

Y3 QQ-plot of the childhood peripheral blood probe-level meta-analysis results: Internalising problems age 7 405

Appendix Z 406

Z1 Leave-one-out plot of the cord-blood meta-analysis age 7 (Cg26668632) 406

Z2 Leave-one-out plot of the childhood peripheral blood meta-analysis age 7 (Cg08884410) 406

Z3 Leave-one-out plot of the childhood peripheral blood meta-analysis age 7 (Cg07283896) 407

Appendix AA 408

AA1 Forest plot of associations between cord blood BCELL cell proportion types and offspring internalising problems at age 3 408

AA2 Forest plot of associations between cord blood CD8T cell proportion types and offspring internalising problems at age 3 408

AA3 Forest plot of associations between cord blood NK cell proportion types and offspring internalising problems at age 3 409

AA4 Forest plot of associations between cord blood CD4T cell proportion types and offspring internalising problems at age 3 409

AA5 Forest plot of associations between cord blood CD14 cell proportion types and offspring internalising problems at age 3 410

AA6 Forest plot of associations between cord blood GRAN cell proportion types and offspring internalising problems at age 3 410

AA7 Forest plot of associations between cord blood BCELL cell proportion types and offspring internalising problems at age 7 411

AA8 Forest plot of associations between cord blood CD14 cell proportion types and offspring internalising problems at age 7 411

AA9 Forest plot of associations between cord blood CD4T cell proportion types and offspring internalising problems at age 7 412

AA10 Forest plot of associations between cord blood CD8T cell proportion types and offspring internalising problems at age 7 412

AA11 Forest plot of associations between cord blood GRAN cell proportion types and offspring internalising problems at age 7 413

AA12 Forest plot of associations between cord blood NK cell proportion types and offspring internalising problems at age 7 413

AA13 Forest plot of associations between peripheral blood BCELL cell proportion types and offspring internalising problems at the age of 7 414

AA14 Forest plot of associations between peripheral blood CD4T cell proportion types and offspring internalising problems at age 7 414

AA15 Forest plot of associations between peripheral blood CD8T cell proportion types and offspring internalising problems at age 7 415

AA16 Forest plot of associations between peripheral blood GRAN cell proportion types and offspring internalising problems at age 7... 415

AA17 Forest plot of associations between peripheral blood MONO cell proportion types and offspring internalising problems at age 7... 416

AA18 Forest plot of associations between peripheral blood NK cell proportion types and offspring internalising problems at age 7 416

Appendix BB.. 417

BB1 Candidate gene-wide analyses 417

Appendix CC 418

CC1 Summary of results of top CpG sites from the probe-level analysis..................... 418

List of Abbreviations

ADHD	Attention deficit hyperactivity disorder
ALSPAC	Avon Longitudinal Study of Parents and Children
ARIES	Accessible Resource for Integrated Epigenomic Studies
BDNF	Brain-derived neurotropic factor
BiB	Born in Bradford
BMI	Body mass index
CBCL	Child behaviour checklist
Chr	Chromosome
CpG	Cytosine-phosphate-guanine
DMR	Differentially methylated regions
DNA	Deoxyribonucleic acid
DOHaD	Developmental Origins of Health and Disease
EAGLE	EArly Genetics and Lifecourse Epidemiology
EDEN	Etude des Déterminants pré et post natals du développement et de la santé de l′Enfant
EFSA	European Food Safety Authority
EWAS	Epigenome-wide association studies
FDR	False discovery rate
FFQ	Food frequency questionnaires
FKBP5	FKBP Prolyl Isomerase 5
GAD	Generalised anxiety disorder
GCSE	General Certificate of Secondary Education
Generation R	Generation Rotterdam
GO	Gene ontology
GWAS	Genome-wide association studies
HPA-axis	Hypothalamic-pituitary-adrenal-axis

INMA	INfancia y Medio Ambiente
IQ	Intelligence quotient
IQR	Inter-quartile range
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LD	Linkage disequilibrium
Mb	Mega base
MoBa	Norwegian Mother and Child Cohort
mQTL	Methylation quantitative trait locus
MR	Mendelian randomisation
nAChRs	Nicotinic acetylcholine receptors
NHS	National Health Service
NR3C1	Nuclear receptor subfamily 3 group C member 1
ODD	Oppositional defiant disorder
PACE	Pregnancy And Childhood Epigenetics
PCA	Principal component analysis
PheWAS	Phenome-wide association study
P-factor	General psychopathology factor
PRS	Polygenic risk score
QQ-plot	Quantile-quantile plot
RCTs	Randomised controlled trials
SD	Standard deviation
SDQ	Strengths and Difficulties Questionnaire
SE	Standard error
SEP	Socio-economic position
SLC6A4	Serotonin transporter gene
SNP	Single nucleotide polymorphism
SVA	Surrogate variable analysis
UK	United Kingdom
WHO	World Health Organization

Chapter 1 – Introduction

1.1 Chapter overview

In this chapter, I review studies that examine the hypothesis of a potential causal effect of smoking and caffeine consumption during pregnancy on offspring mental health outcomes, summarizing research applying different study designs, including observational, genetic, and epigenetic studies. Next, I outline the theory that motivated the investigation of offspring DNA methylation as a potential mediator of the associations under investigation. Lastly, I describe the limitations of the current evidence and explain how the investigation of a molecular mechanism linking intrauterine smoking and caffeine exposure to offspring mental health outcomes could add to the causal evidence base of this relationship.

1.2 Developmental Origins of Health and Disease hypothesis and offspring mental health outcomes

The Developmental Origins of Health and Disease (DOHaD) hypothesis focusses on investigating how the quality of the early environment (pre- and post-natal) can shape offspring's susceptibility to health problems later in life (Wadhwa et al., 2009). One famous example of DOHaD research is the investigation of the effects of the Dutch famine at the end of World War II in the Netherlands on offspring's later health outcomes (Lumey et al., 2007). Data on this event allowed researchers to study the effects of deprivation of crucial nutrients and exposure to excessive stress during pregnancy. Results of this research indicated that malnutrition and stress during pregnancy led to increased risk for numerous diseases in offspring, including mental health problems, such as schizophrenia and affective disorders (Hoek et al., 1998; Lumey et al., 2007). Ever since, the interest in the effects of exposures early in development has rapidly increased, and there is a large and growing body of work in the fields of psychology and psychiatry investigating the effect of exposures during pregnancy on offspring mental health outcomes (Abdul-Hussein et al., 2020; Gage et al., 2016; Lieshout & Krzeczkowski, 2016). These studies are based on the hypothesis that the quality of the prenatal environment "programmes" offspring's susceptibility to mental health problems (Lieshout & Krzeczkowski, 2016; O'Donnell & Meaney, 2017). In line with this hypothesis, this thesis explores whether maternal smoking and caffeine

consumption during pregnancy have a causal effect on offspring's mental health outcomes (potential mechanisms are outlined in section 1.7).

1.3 Mental health problems in childhood and adolescence

Child mental health problems are commonly divided into externalising and internalising behaviour problems (Achenbach et al., 2016; Angold et al., 1999). The former describes mental health problems that manifest through the child expressing behaviours that act on the external environment and incorporates mental illnesses that describe behaviours such as impulsivity, aggression, and substance use. The latter describes mental health problems that include behaviours that are affecting the child's inner state rather than the outer environment, such as withdrawal, anxiety, and depression symptoms. The concepts of externalising and internalising problems are commonly applied in research as well as clinical practice of child psychology and psychiatry (Achenbach et al., 2016). While externalising and internalising problems have been found to capture two different constructs (Achenbach, 1966), they are not mutually exclusive, might change over the course of development, and often co-occur (Angold et al., 1999; Bubier & Drabick, 2009; Sallis et al., 2019). Recent research has even proposed that the high number of comorbidities of diagnoses of mental health problems can be summarised into one common general psychopathology factor (or P-factor). The P-factor combines internalising and externalising problems and focuses on understanding the joint impact of both externalising and internalising problems in children (Murray et al., 2016; Sallis et al., 2019). Shared genetic factors influencing many mental health problems have been proposed as an explanation for the P-factor (Lahey et al., 2011). However, even after accounting for the common variance of externalising and internalising problems, a significant amount of variance unique to externalising and internalising problems remains (Murray et al., 2016). A study combining data from three large European cohort studies has found that in terms of quality of life later in life, the P-factor, combining shared variance of internalising and externalising problems during childhood, was the strongest predictor for negative outcomes. However, the unique variance of externalising problems was not associated with negative outcomes later in life. Despite this, the unique variance of internalising problems was still predictive of negative outcomes, including depression, anxiety, and

wellbeing in early adulthood (Sallis et al., 2019). Internalising problems are more difficult to assess, especially in young children (Goodman & Scott, 1999; Maurice-Stam et al., 2018). Both externalising and internalising problems in children are most commonly assessed through interviews or parent-based self-report questionnaires. Whereas externalising behaviours can be observed by external reporters (e.g., teachers), internalising behaviour problems are less evident to external reporters, and require introspective skills of the individual to report. In addition, parents' own mental health problems have been found to impact reporting of their children's mental health outcomes (Gartstein et al., 2009; Najman et al., 2001) and the importance of using multiple informants on child internalising and externalising problems has been stressed to remove unique variance of the type of informant (Sallis et al., 2019). In sum, both externalising and internalising problems in childhood, specifically if co-occurring, are associated with negative health outcomes later in life, stressing the need to detect early manifestations of externalising and internalising problems and their determinants.

1.4 Prevalence of offspring externalising and internalising problems

Mental health problems in young people are significantly contributing to the global burden of health and disease (Erskine et al., 2015). It is estimated that about 50% of mental health problems have their origin in childhood (under the age of 14), but are often not treated until adulthood (Dick & Ferguson, 2015; Patel et al., 2007). In the United Kingdom (UK), national representative surveys were conducted to estimate the prevalence of mental health problems amongst young people aged 5 to 19 years. Mental health diagnoses of young people were derived according to diagnostic criteria of the International Classification of Disease (World Health Organization, 1992) and the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013) in the years 1999 (N = 10,438; obtained from Child Benefit records) (Meltzer et al., 2000), 2004 (N = 12,294; data obtained from Child Benefit records) (Green et al., 2005) and 2017 (N = 18,029 children with data in National Health Service (NHS) Patient register) (Sadler et al., 2018). The comparison of prevalence rates of these three reports showed a slight increase in the diagnosis of any mental health problem between

the ages 5 and 15, with a prevalence of 10% in 1999, 10% in 2004 and 11% in 2017 (Figure 1.1) (Sadler et al., 2018).

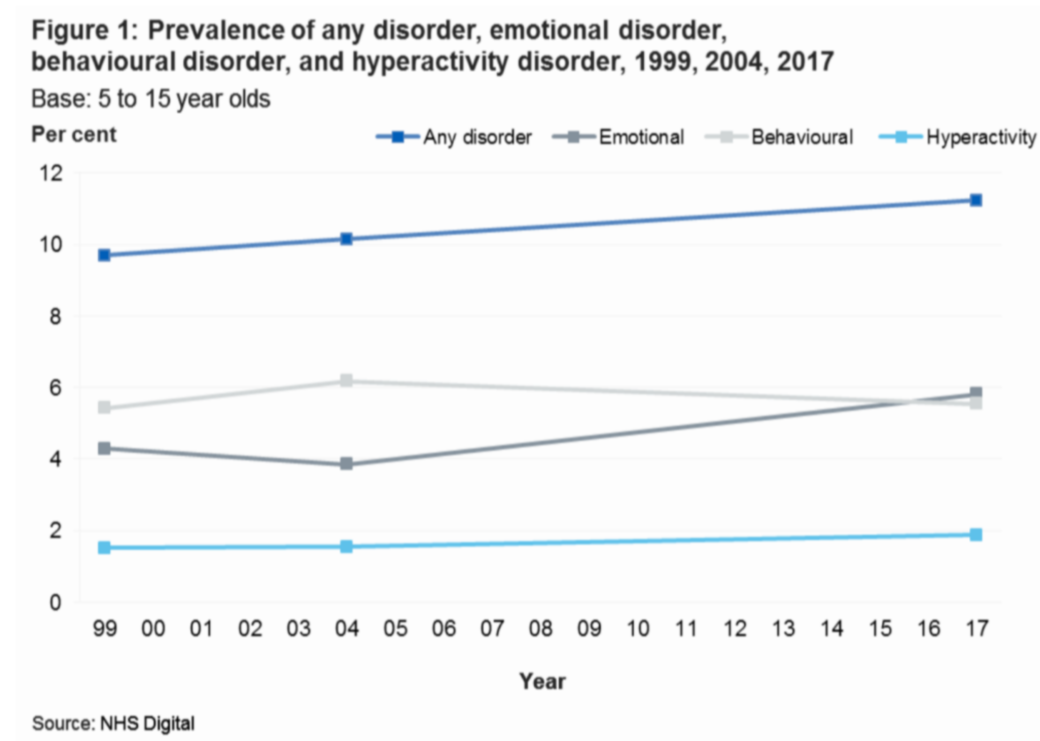


Figure 1.1 Prevalence rates of mental health problems in young people from 1999 to 2017 *(credit: Sadler et al., 2018).*

As shown in Figure 1.1, whereas externalising behaviour problems remained relatively stable over time, internalising behaviour problems increased from approximately 4% in 1999 and 2004, to approximately 6% in 2017 (Sadler et al., 2018). The most recent survey found that in 2017, 1 out of 12 children (age range 5 to 19) in England suffered from an internalising disorder, such as anxiety and depression symptoms. Prevalence rates for internalising disorders were higher for children identified as girls (10%) than children identified as boys (6%). In contrast, externalising problems were more commonly diagnosed in children identified as boys than girls. One out of 20 children met the diagnosis for a conduct disorder (prevalence: boys = 6%; girls = 3%) and 1 out of 6 children met the diagnosis for a hyperactivity disorder (prevalence: boys = 3%; girls = 1%) (Sadler et al., 2018). Overall, childhood mental health problems appear to have slightly increased over the past 20 years (Sadler et al., 2018) and prevention

strategies have been proposed to be the most efficient strategy to decrease mental health problems in children (Erskine et al., 2015). Successful prevention of mental health problems in children requires the identification of modifiable and quantifiable risk factors early in development (Erskine et al., 2015). This thesis will investigate the potential of smoking and caffeine consumption during pregnancy as a modifiable target for preventing offspring mental health disorders.

1.5 Prevalence of smoking and caffeine consumption during pregnancy

Before discussing the relationship of maternal smoking and caffeine consumption on offspring mental health problems more closely, I am going to elaborate on the relevance to public health by highlighting the prevalence of smoking and caffeine consumption during pregnancy with a particular focus on the UK and other European countries. Public awareness of the harmful effects of smoking during pregnancy has been increasing over the last 40 years and European strategies have been developed to promote guidelines for smoking cessation during pregnancy (Tobacco Free Initiative (World Health Organization), 2013). However, any effect of caffeine is less well understood, and most pregnancy guidelines recommend limiting caffeine consumption during pregnancy (discussed in more detail in section 1.9.1) (Reyes & Cornelis, 2018). Given the prevalence of smoking and caffeine consumption, an effect on child mental health, however small, would be important to note and may inform prevention and intervention strategies for smoking and caffeine use during pregnancy and offspring mental health problems.

1.5.1 Smoking

Despite established detrimental effects of smoking during pregnancy on offspring health outcomes, including low birth weight, prematurity and stillbirth, and guidelines recommending smoking cessation during pregnancy, this behaviour is still quite prevalent in Europe. Stopping smoking early during pregnancy has been found to substantially reduce the risk for offspring health outcomes (Diamanti et al., 2019, 2019; Giglia et al., 2006; Robinson et al., 2010; Smedberg et al., 2014). Even though most smoking women want to stop smoking during pregnancy, only around 30% of mothers who have smoked prior to pregnancy manage to stop smoking during pregnancy (Giglia et al., 2006; Robinson et al., 2010).

Chapter 1 – Introduction

An extensive meta-analysis reviewed 295 studies that have reported the prevalence of smoking during pregnancy across 43 countries over the past 30 years (Lange et al., 2018). Studies were included if primary data was available, samples contained lifetime non-smokers, and they were representative of the general population of the corresponding country (national/regional representative surveys or studies based on hospital data with complete reporting of births within a specific timeframe). Countries for which multiple studies were found, estimates were meta-analysed using random effects. Counties for which merely one or no study was available, regression models were used to predict the prevalence of smoking during pregnancy using country-specific indicators (e.g., country gross domestic product, gender inequality-index, prevalence of women smoking outside of pregnancy, etc.). These county estimates were then meta-analysed to estimate global and World Health Organization (WHO) region prevalence of smoking during pregnancy. Whereas the prevalence of smoking during pregnancy globally decreased over time and was estimated to be less than 2%, the prevalence in European regions was highest, with 8% of women smoking during pregnancy. The study estimated that globally, amongst women who report smoking during pregnancy, more than 70% of women do so daily. Overall, there is evidence that women who smoke during pregnancy substantially reduce the number of cigarettes smoked per day (Ershoff, 2000), with most women smoking fewer than 10 cigarettes per smoking day (Lange et al., 2018; Smedberg et al., 2014). In fact, for the European region it was estimated that only 30% of women who smoked daily before becoming pregnant continued to smoke daily during pregnancy.

Prevalence rates of smoking during pregnancy vary substantially between and within countries and regions (Lange et al., 2018), reflecting the complexity of this behaviour and how embedded it is in socioeconomic structures. For instance, Lange and colleagues (Lange et al., 2018) report that within the UK on average 23% of women smoke at some point during pregnancy (N studies = 19, range sample size: 82-16,865) and that the meta-analysed UK studies show substantial heterogeneity ($I^2 = 96\%$, P-value < 0.0001) (Lange et al., 2018). Prevalence rates across UK studies ranged between 14% in a study assessing smoking during pregnancy between 2007 to 2010 (Cooper et al., 2013), to 40%, in a study that

assessed smoking during pregnancy between 1983 to 2003 (Delpisheh et al., 2007).

1.5.2 Caffeine

The effects of intrauterine caffeine exposure on offspring health outcomes are not well understood yet, which is reflected in lack of evidence-based pregnancy guidelines for caffeine use during pregnancy (Reyes & Cornelis, 2018). There are no nationally representative studies available for caffeine consumption during pregnancy (Verster & Koenig, 2018). Prevalence rates from prospective cohort studies suggests that the majority of women (96%) reduce or stop caffeine consumption during pregnancy, and 65% show caffeine aversion in the beginning of pregnancy (Lawson et al., 2002, 2004). Women that reduce their caffeine consumption during pregnancy consume on average only half the amount of caffeine that they consumed before becoming pregnant (Knight et al., 2005; Lawson et al., 2004). However, prospective cohort studies report that 60% to 75% of women continue consuming caffeine during pregnancy, with around 15% consuming more than 200 mg of caffeine a day (Chen et al., 2014; Reyes & Cornelis, 2018; Weng et al., 2008). Caffeine consumption is culturally bound, and the sources of caffeine may vary between different countries and be differentially socially patterned (Miyake et al., 2019; Reyes & Cornelis, 2018; Treur et al., 2016). For instance, the most commonly consumed source of caffeine in Europe is coffee, except for the United Kingdom (UK), where consumption of caffeinated black tea is more common than coffee (Reyes & Cornelis, 2018; Treur et al., 2016). In Japan and China, the most common source of caffeine is green tea and in South America it is Yerba Mate (Reyes & Cornelis, 2018; Treur et al., 2016). Whereas drinks containing natural sources of caffeine were found to be culturally specific, caffeinated soda and energy drinks were found to be similarly common across cultures (Reyes & Cornelis, 2018). A study comparing the relationship between smoking and caffeine consumption using data from prospective birth cohorts from the Netherlands and the UK found that caffeinated tea consumption in the Netherlands was associated with higher education and not associated with smoking initiation or smoking persistence. Yet in the UK, caffeinated tea consumption was not associated with education or socio-economic position (SEP), but with smoking initiation and persistence (Treur et al., 2016). This

indicates how culturally bound caffeine consumption is and that consumption can be differentially socially patterned even within countries that are perceived to be culturally comparable.

1.6 Observational evidence of prenatal smoking and caffeine exposure and mental health outcomes

1.6.1 Smoking

Several observational epidemiological studies have investigated the relationship between smoking during pregnancy on offspring mental health outcomes (e.g., see review by Tiesler & Heinrich, 2014). Most of this research stems from longitudinal birth cohort studies, which are a useful resource to study the impact of early exposures on later offspring health outcomes, while allowing to control for potential confounding variables, eliminating reverse causation and recall bias as potential alternative explanations for any observed association. After randomised controlled trials (RCTs), which are often unethical and/or unfeasible for investigating intrauterine effects, they are the best source of evidence for interrogating these causal questions (Davey Smith, 2008).

Data from the Finnish medical birth register, comprising a large representative sample of around 180,000 mothers and their children, found that the prevalence rates for offspring obtaining any psychiatric diagnosis between the ages 9 to 20 years increased in a dose-response manner with the average number of cigarettes smoked per day during pregnancy (Ekblad et al., 2010). The risk of having any psychiatric disorder was 14% for offspring whose mothers did not report smoking during pregnancy, 21% for offspring whose mothers smoked 10 or fewer cigarettes per day, and 25% for offspring whose mothers smoked more than 10 cigarettes per day.

In contrast to the number of studies that investigated the relationship of maternal smoking during pregnancy and offspring externalising problems, comparatively few have investigated its association with offspring internalising problems (Dolan et al., 2016; Easey & Sharp, 2021; Tiesler & Heinrich, 2014). A systematic review investigating prenatal and early postnatal risk factors for externalising behaviour disorders identified smoking during pregnancy as one of the most

researched prenatal risk factors. Out of 11 studies that investigated smoking during pregnancy on offspring externalising problems, 8 reported evidence for an association with offspring attention deficit hyperactivity disorder (ADHD) symptoms (Latimer et al., 2012). Another review that looked at externalising and internalising problems in children and intrauterine smoke exposure also reported evidence for an association between smoking during pregnancy and offspring externalising problems, but there was not sufficient evidence to confirm a causal effect (Tiesler & Heinrich, 2014). However, the number of studies investigating the association of smoking during pregnancy and internalising problems in offspring was underrepresented and mixed findings were reported in those studies inspecting an effect on internalising problems (Tiesler & Heinrich, 2014). A recent meta-analysis of eight cohorts and two case control studies found evidence for an association between smoking during pregnancy and an increased risk of offspring mood disorders (bipolar disorder: RR = 1.44, 95% CI = 1.15 to 1.80; and depression: RR = 1.44, 95% CI = 1.21 to 1.71) (Duko et al., 2020). The forest plot of this meta-analysis is displayed in Figure 1.2.

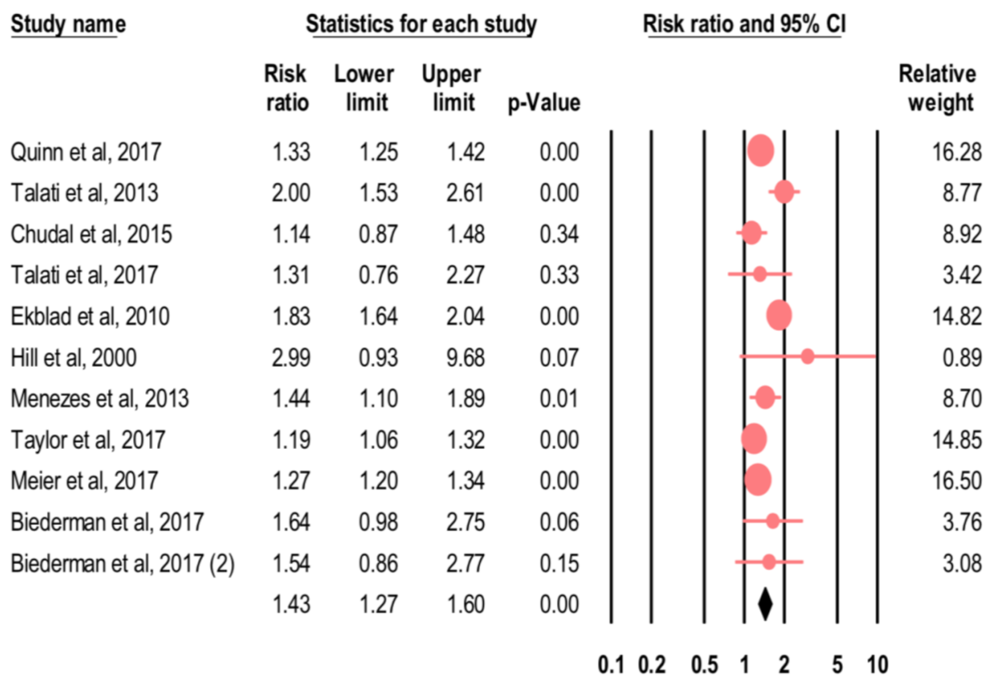


Figure 1.2. Forest plot of the studies included in the meta-analysis by Duko et al. investigating prenatal smoking and the risk for offspring mood disorders *(credit: Duko et al., 2020).*

Whereas the meta-analysis by Duko and colleagues (Duko et al., 2020) focussed on diagnoses of mood disorders (age range 8 to 41 years) instead of behavioural manifestations of internalising problems, other studies have specifically inspected the association between prenatal smoking and risk for childhood internalising problems. A study using the Norwegian Mother and Child Cohort (MoBa; N = 90,040 mother-child pairs) found that early smoking during pregnancy increases the risk for internalising problems in toddlers (1.5 to 3 years) in a dose-response manner, even when controlling for important confounding variables (paternal smoking, maternal education, maternal age, maternal depressive and anxiety symptoms, maternal alcohol consumption, parity, gestational age at birth, and smoking in previous pregnancies) (Moylan et al., 2015). This is in line with some previous research (Ashford et al., 2008b; Ekblad et al., 2010; Indredavik et al., 2007; Menezes et al., 2013), but in contrast to other studies which did not find evidence for an association between maternal smoking and offspring internalising problems (Brion et al., 2010; Höök et al., 2006; Lavigne et al., 2011; Roza et al., 2009).

1.6.2 Caffeine

Few research studies have investigated the effects of caffeine consumption during pregnancy on offspring externalising and internalising problems, generally producing mixed results. Whereas some studies using data from prospective birth cohorts report no association with offspring outcomes after adjusting for covariates (Berglundh et al., 2020a; Del-Ponte et al., 2016; Linnet et al., 2009; Loomans et al., 2012; Miyake et al., 2019), those that found evidence for associations are inconsistent about effects of different timings of exposure during pregnancy, types of caffeine source, and type and onset of offspring problem behaviours (Bekkhus et al., 2010; Hvolgaard Mikkelsen et al., 2017; Klebanoff & Keim, 2015). For instance, data from a large Danish birth cohort study (N = 47,491) reported evidence for an increased risk for offspring mental health problems at age 11 in association with high consumption of caffeinated tea and coffee (> 8 cups per day) during the second but not during the third trimester of pregnancy (15 weeks gestation) (Hvolgaard Mikkelsen et al., 2017). Models were adjusted for a variety of significant covariates, such as maternal age, body mass index (BMI), and smoking during pregnancy, as well as mutually adjusted for

either coffee or tea consumption. High consumption of caffeinated coffee was associated with increased risk for offspring externalising disorders (oppositional-conduct disorders, hyperactivity-inattention disorders), whereas high consumption of caffeinated tea was associated with increased risk for offspring internalising disorders (anxiety-depressive disorders) but only oppositional-conduct disorders from the externalising disorders group (Hvolgaard Mikkelsen et al., 2017). Both high consumption of caffeinated tea and coffee were associated with increased risk for any psychiatric disorder. None of these associations were replicated with maternal caffeine consumption during the third trimester of pregnancy. When considering total caffeine consumption during the second and third trimester of pregnancy (based on caffeine content of one cup of coffee = 100 mg and one cup of tea = 50 mg), a significantly reduced risk for offspring oppositional-conduct disorder, hyperactivity-inattention disorder, and any psychiatric disorder was observed for low caffeine consumption (Figure 1.3).

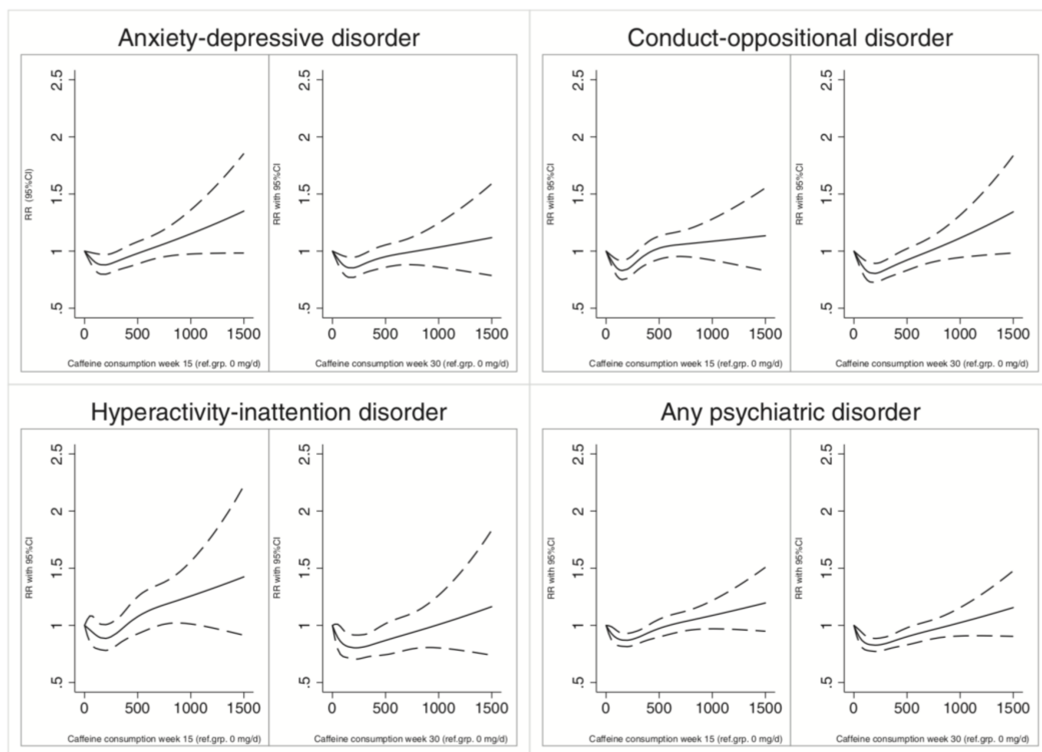


Figure 1.3. Risk for offspring mental health problems at the age of 11 in association with total maternal caffeine consumption at 15- and 30-weeks gestation *(credit: Hvolgaard Mikkelsen et al., 2017).*

The authors hypothesised that the slightly increased risk for offspring mental health problems of mothers who did not consume coffee or tea, compared to offspring of mothers who did consume some coffee or tea, may be explained by the former being more likely to consume caffeine from other sources, such as caffeinated soda and energy drinks. They also hypothesised that exposure during the second trimester may have a stronger effect on offspring's brain development and therefore explain why only second trimester but not third trimester exposure to caffeine was associated with offspring behavioural outcomes. Despite acknowledging the potential for confounding through genetics or socioeconomic factors, the authors concluded that mothers should avoid drinking high amounts of caffeine during pregnancy in order to avoid behavioural problems in their offspring.

In contrast, a smaller study from the US using data from around 2,000 mother-child pairs found evidence for associations between serum paraxanthine, a biomarker for maternal caffeine consumption, and offspring mental health problems, only in the third trimester but not earlier in pregnancy (Klebanoff & Keim, 2015). Paraxanthine is the main metabolite of caffeine, which shows a similar half-life (3.5 to 5 hours), captures caffeine consumption and metabolism, and shows more stability over the day than serum caffeine (Klebanoff & Keim, 2015). The study observed inconsistent associations between paraxanthine under 20 weeks gestation and offspring mental health outcomes, with an increased risk for internalising problems at the age of 4 but an opposite direction of effect for internalising problems at the age of 7 (although not statistically significant). Likewise, statistical evidence for associations was found for externalising problems at the age of 7 but not at the age of 4. Due to these inconsistent effects and lack of representation of very high caffeine consumption in their sample, the authors concluded low to moderate amounts of caffeine during pregnancy to not be hazardous for offspring mental health outcomes (Klebanoff & Keim, 2015). Similarly, inconsistent results were reported by the Norwegian birth cohort study MoBa (N = 32,927), that reported evidence for associations with some offspring ADHD symptoms only for caffeinated soft drink but not coffee or tea consumption during pregnancy (Bekkhus et al., 2010). The associations with caffeinated soft drink consumption were only consistent across the second and

third trimester for offspring hyperactivity but not for inattention symptoms (Bekkhus et al., 2010).

In sum, observational evidence for an effect of intrauterine exposure to smoking and caffeine on offspring mental health outcomes shows mixed findings and provides confusing recommendations for intake during pregnancy. It should be noted that I could not find any systematic reviews of the effects of prenatal caffeine exposure on offspring mental health problems. The abovementioned studies that found evidence for an association with offspring mental health problems have an increased likelihood of being false-positives, due to high number of tests that were run to investigate different timings of exposure and effects on various offspring mental health outcomes.

1.7 Biological effects of intrauterine exposure to smoking and caffeine on offspring mental health development

The observational associations between smoking and caffeine consumption during pregnancy could be biologically plausible, as is suggested by human and animal studies that investigated the chemical compounds of tobacco and caffeine and their potential effect on offspring's neurodevelopment. Chemicals contained in tobacco and caffeine can cross the placental barrier and pregnancy related changes in metabolism of these chemicals further increase the risk for an effect on offspring's neurodevelopment. The next section elaborates on the potential mechanisms underlying the association of smoking and caffeine consumption during pregnancy with offspring's mental health development.

1.7.1 Biological pathways of smoking and caffeine during pregnancy on offspring's brain development

There are several potential mechanisms by which intrauterine exposure to smoking might "programme" offspring's physiological development and therewith increase the risk for offspring's mental health outcomes (Lieshout & Krzeczkowski, 2016). Cigarettes contain over 4,000 different chemicals that could individually or interactively exert effects on offspring (Abbott & Winzer-Serhan, 2012; McDonnell & Regan, 2019). In animal and human studies that investigate the effects of smoking on mental health and behavioural problems, nicotine is the most researched chemical (Abbott & Winzer-Serhan, 2012). A potential pathway

for how intrauterine exposure to nicotine links to behavioural problems might be through the activation of nicotinic acetylcholine receptors (nAChRs), which can influence the development of the nervous system (Abbott & Winzer-Serhan, 2012; Smith et al., 2010). The effects of nAChRs activation on the developing nervous system are based on animal studies, with the majority studying effects in rodents, but the translation of effects to human development is difficult (Garner, 2014; Keeler & Robbins, 2011; Kenney & Müller, 2017). Furthermore, animal studies suggest that intrauterine nicotine exposure might exert effects on offspring’s monoamine system (Blood-Siegfried & Rende, 2010; Xu et al., 2001), which includes the regulation of the neurotransmitters dopamine, noradrenaline, and serotonin, that are linked to mood and anxiety disorders (Goddard et al., 2010) conduct disorder (Baler et al., 2008), as well as aggressive and substance use behaviours (Seo et al., 2008). This is supported by further animal experiments, which found that prenatal exposure to nicotine led to more anxiety type of behaviours in rodents (Abbott & Winzer-Serhan, 2012).

In comparison to smoking, the potential teratogenic effects of caffeine on offspring mental health and behavioural outcomes are less well studied (Ross et al., 2015). Epidemiological studies of the effects of caffeine on health outcomes have predominantly studied the effects of caffeine through assessment of consumption of caffeinated coffee, neglecting that coffee contains over 1,000 other compounds, as well as the effects of other sources of caffeine (van Dam et al., 2020). Similar as for smoking, most mechanistic studies investigating prenatal effects of caffeine on offspring health outcomes rely on animal studies (Qian et al., 2020). In mice, prenatal caffeine exposure has been associated with accumulation of caffeine in offspring’s brain and changes in neurodevelopment, such as persistent changes to the development of the hippocampus (Silva et al., 2013). In humans, hippocampal differences have been found to be associated with psychiatric disorders, such as major depressive and bipolar disorders, and schizophrenia (Sala et al., 2004), as well as with internalising problems in children and adolescents (Andre et al., 2020; Koolschijn et al., 2013). Potential effects of caffeine on neurodevelopment include delayed neuronal migration and inflated neuronal excitability likely due to caffeine acting on the adenosine receptor (Fuzik et al., 2019; Silva et al., 2013). Evidence for an effect of caffeine

exposure during pregnancy on human neurodevelopment is limited. There is some evidence from human studies that report positive effects of caffeine on neurodevelopmental outcomes in infants (Christensen et al., 2021). For instance, caffeine is commonly applied as treatment of apnea in preterm infants because of its potential to act as an adenosine receptor antagonist and its beneficial effects on respiratory and neurological functioning in those children (Abdel-Hady et al., 2015; Kua & Lee, 2017). Through caffeine's influence on the adenosine receptor, it may change the regulation of neurotransmitters such as dopamine, noradrenaline, and serotonin (Abdel-Hady et al., 2015; Fredholm et al., 1999), and exert effects on mental health symptoms of depression and anxiety (Calker et al., 2019).

However, it remains unclear whether the neurological effects of prenatal smoking and caffeine that have been observed in animal studies translate to prenatal effects in humans, and if they do, whether they can explain the increased risk for later behavioural and mental health problems that have been associated with prenatal exposure to smoking and caffeine.

A recent neuroimaging study conducted in the United States investigated the relationship between caffeine exposure during pregnancy, offspring psychopathology (including externalising and internalising problems) and neurodevelopmental outcomes in 8- to 9-year-old children (N = 1,200) (Christensen et al., 2021). Maternal caffeine consumption during pregnancy was assessed through maternal self-report 8 to 10 years post-pregnancy and therefore results should be interpreted with caution, as risk for recall bias is high. Evidence of neurological differences (differences in 2 of 27 studied fibre tracts) was found between children exposed and unexposed to caffeine in utero, as well as increased levels of psychopathology amongst exposed versus unexposed children. However, there was no evidence for the neurological differences mediating the relationship of caffeine exposure during pregnancy and childhood psychopathology.

In sum, the above-mentioned literature highlights that smoking and caffeine consumption expose the fetus to various compounds that may biologically impact offspring's neurodevelopment and behavioural outcomes. Though, most of these mechanistic studies rely on results of animal data and it is unclear whether these

results translate to human populations.

1.7.2 Pregnancy related changes in tobacco and caffeine metabolism

In order to understand the biological effects that smoking and caffeine consumption may have on offspring mental health, it is important to consider pregnancy related changes in the metabolism of nicotine and caffeine. Neither the placenta nor the fetus is able to metabolize nicotine or caffeine, leaving the clearance of these substances dependent on maternal metabolism (Bowker et al., 2015; Grosso & Bracken, 2005; McDonnell & Regan, 2019; van Dam et al., 2020; Wickström, 2007; Yu et al., 2016). Whilst there may be individual differences in the metabolism rate of chemicals contained in cigarettes and caffeine, general changes during pregnancy have been observed: nicotine metabolism increases around the second trimester of pregnancy (Bowker et al., 2015), while caffeine metabolism decreases during pregnancy (Yu et al., 2016). Faster nicotine metabolism may cause smoking mothers to smoke more cigarettes to alleviate their nicotine withdrawal symptoms and may be one of the reasons why nicotine replacement therapy, using the same levels as used outside of pregnancy, has been found not to be effective during pregnancy (Bowker et al., 2015; Jauniaux et al., 1999). In contrast, decreased metabolism of caffeine during pregnancy may lead mothers to reduce their caffeine intake naturally. Throughout pregnancy, the metabolism of caffeine gradually decreases and in the second and third trimester of pregnancy the half-life of caffeine can last up until 18 hours, which is about four times longer than outside of pregnancy (Grosso & Bracken, 2005). If mothers – in addition to consuming caffeine – smoke during pregnancy, their caffeine metabolism will increase again (Grosso & Bracken, 2005), which may allow them to uphold their habitual caffeine consumption. However, as most mothers who smoke during pregnancy were already smoking before becoming pregnant, it seems more plausible that mothers who smoke during pregnancy also reduce their caffeine consumption to some degree. Interestingly, a RCT of pregnant women ($N = 207$) consuming high levels of caffeine (> 300 mg/day), who were allocated to either consumption of caffeinated or decaffeinated instant coffee, did not find evidence of a difference in smoking behaviour between the groups (Bech et al., 2007). However, given the small sample size of this trial, this absence of clear evidence could likely be due to insufficient statistical power to detect an effect.

Nevertheless, if this finding replicates in larger samples, it could indicate that changes in the metabolic rate of caffeine during pregnancy do not influence maternal smoking behaviour. Due to the changes of caffeine metabolism during pregnancy and the potential of caffeine readily crossing the placenta barrier, the European Food Safety Authority (EFSA) has alerted *"…unborn children to be the most vulnerable group for adverse effects of caffeine among the general population."* (EFSA Panel on Dietetic Products, Nutrition & Allergies (NDA), 2015, p. 33).

1.7.3 Epigenetics as a pathway linking intrauterine exposure to smoking and caffeine on offspring mental health outcomes

The identification of a molecular pathway that could explain how prenatal smoking and caffeine exposure may become biologically embedded to increase the risk for offspring mental health outcomes could strengthen the causal evidence base. Epigenetics may prove as an intermediate molecular pathway explaining how genes interact with environmental exposures, such as intrauterine smoking and caffeine exposure, to increase offspring's risk for disease (Cortessis et al., 2012; Perera & Herbstman, 2011). Epigenetic changes are changes in gene expression in the absence of modifications to the DNA sequence (Bird, 2007). The most studied epigenetic process is Deoxyribonucleic acid (DNA) methylation, which is mainly associated with gene deactivation through the addition of a methyl group to a cytosine-phosphate-guanine (CpG) island near to a gene promoter site (Bird, 2007). There is evidence that DNA methylation is influenced by genetic as well as environmental factors (Teh et al., 2014). Evidence for this stems from twin studies with genetically identical twins. Whereas methylation patterns of monozygotic twins were found to be similar in early ages, they became more dissimilar as twins grew older, especially when having more divergent lifestyles (Fraga, Ballestar, Paz, Ropero, Setien, Ballestar, Heine-Suñer, Cigudosa, Urioste, & Benitez, 2005). DNA methylation undergoes profound changes during embryonic development and is therefore considered a promising pathway that may link prenatal exposures to offspring health outcomes later in life (Felix et al., 2018). The effect of prenatal exposures on offspring DNA methylation can be tested through epigenome-wide association studies (EWAS). EWAS are considered a hypothesis-free approach, which tests associations with a

phenotype of interest and differing levels of DNA methylation at individual CpG sites.

This thesis focusses on the potential effect of maternal smoking and caffeine consumption during pregnancy on offspring DNA methylation and investigates whether these exposure-associated DNA methylation changes may increase the risk for offspring mental health problems. The next section reviews previous EWAS that have been conducted in humans to assess associations with prenatal smoking and caffeine exposure, as well as offspring mental health outcomes.

1.7.3.1 EWAS of smoking and caffeine consumption during pregnancy

Smoking during pregnancy has been found to be robustly associated with offspring DNA methylation assessed in offspring's cord blood. In a large EWAS meta-analysis of 13 birth cohorts (N = 6,685), differing DNA methylation levels were detected at more than 6,000 CpG sites (Joubert et al., 2016). These DNA methylation differences were mostly maintained until childhood (Joubert et al., 2016). Despite the high phenotypic correlation between smoking and caffeine consumption during pregnancy, no EWAS has investigated the effect of maternal caffeine consumption on offspring cord blood DNA methylation yet.

1.7.3.2 EWAS of childhood mental health problems

Studies investigating DNA methylation and childhood mental health problems have predominantly focussed on DNA methylation at specific gene regions that were selected based on an *a priori* hypothesis (Barker, Walton, & Cecil, 2018; Jones et al., 2018). Often, hypotheses of these candidate gene studies are based on DNA methylation differences found in animal studies (Barker, Walton, & Cecil, 2018; Jones et al., 2018). However, for understanding the epigenetic variation associated with complex phenotypes, such as childhood mental health problems, which are influenced by the interplay of multiple genes, a hypothesis generating approach appears to be more appropriate. The lack of overlap between candidate gene and epigenome-wide DNA methylation studies of mental health phenotypes highlights the limited knowledge that we currently have about the molecular underpinnings of mental health problems in humans (Barker, Walton, & Cecil, 2018). Whereas many EWAS have investigated the association between risk factors and offspring DNA methylation, only few studies have investigated the

association between DNA methylation and childhood mental health outcomes (Barker, Walton, & Cecil, 2018). A small EWAS investigated the association between DNA methylation in saliva of 190 children, aged 5 to 14 years, discordant for maltreatment. Whereas DNA methylation at three CpG sites was associated with depressive symptoms, no differences between DNA methylation levels of maltreated and non-maltreated children were observed (Weder et al., 2014). Further, a small study of 18 monozygotic twins discordant for adolescent depression, found increased DNA methylation at one CpG site to be associated with depression, which was replicated in post-mortem brain tissue of veterans discordant for depression (Dempster et al., 2014). No overlap in results was observed between these two studies, neither between the CpG sites nor their annotated genes. Similar to the pattern found in observational studies, in comparison to internalising problems, more EWAS of externalising problems have been conducted (Barker, Walton, & Cecil, 2018). At birth, DNA methylation differences at 7 CpG sites were found to differentiate between young children's conduct symptoms (N = 260) (Cecil et al., 2018), 13 CpG sites to be associated with childhood ADHD (N = 892) (Walton et al., 2017) and 30 CpG sites with oppositional defiant disorder trajectories (N = 624) (Barker, Walton, Cecil, et al., 2018). Yet again, little overlap was observed between these differentially methylated CpG sites (Barker, Walton, Cecil, et al., 2018). Noteworthy, all these EWAS have been conducted in the same UK based birth cohort and thus uncertainty remains whether these findings replicate in other cohorts. A recent large, international meta-analysis of DNA methylation and childhood ADHD (N = 2,374) has found DNA methylation at 9 CpG sites at birth to be associated with ADHD at age 7-11 years (Neumann et al., 2020). None of the CpG sites reported in the previous ADHD EWAS by Walton and colleagues (Walton et al., 2017) replicated in the study by Neumann and colleagues (Neumann et al., 2020).

1.8 Problems with observational evidence

The inconsistent evidence for associations between maternal smoking and caffeine consumption and offspring mental health outcomes is likely due to the majority of research being reliant on observational designs and the heterogeneous nature of these designs and populations studied. Due to the established adverse effects of smoking and the uncertainty about the effects of caffeine during

pregnancy, it is not possible to conduct conventional controlled trials that randomly allocate mothers to either consuming or abstaining from smoking or caffeine during pregnancy. Randomised controlled trials (RCTs) of smoking or caffeine cessation interventions during pregnancy could theoretically be used to make causal inferences about potential intrauterine effects on offspring mental health outcomes. RCTs of smoking cessation interventions during pregnancy that have investigated effects on offspring health outcomes have predominantly focussed on birth outcomes, such as low birth weight and preterm birth (Chamberlain et al., 2017; Gresham et al., 2014; Lumley et al., 2009). Very few RCTs have investigated outcomes later in development (Coleman et al., 2012, 2015) and I am not aware of any that assessed effects on offspring mental health or behavioural outcomes. This may be because of limited statistical power of such studies to detect small to moderate effects and the risk of bias because of low success rates of cessation interventions, high dropout rates, small sample sizes, and confounding by postnatal smoking relapse. Further, withdrawal symptoms and problems with blinding treatment and control conditions (e.g., having an appropriate placebo) complicate the conduct and interpretation of RCTs (Bech et al., 2007; Chamberlain et al., 2017).

Most of the data that investigates prenatal effects stems from prospective birth cohort studies, as they allow ruling out reverse causation because of the known temporal sequence of events (it is not possible that offspring mental health problems have caused mothers to smoke or consume caffeine during pregnancy). However, other problems of observational research remain. The next section outlines the most important biases of observational associations of maternal smoking and caffeine consumption during pregnancy and offspring mental health outcomes in prospective birth cohort studies.

1.8.1 Selection bias

Two main types of missing data problems of longitudinal birth cohort studies contribute towards selection bias. First, longitudinal studies miss data of people who do not enrol in the study in the first place and therewith also the representation of associated characteristics of people who chose not to participate (or have not been approached to participate). Second, longitudinal studies miss data of people who initially participated but drop out of the study later on

(Hammerton & Munafò, 2021). Both of these missing data mechanisms are themselves associated with mental health problems, as research has shown that people with an increased risk for mental health problems are less likely to enrol and less likely to stay within cohort studies (Martin et al., 2016; Taylor, Jones, et al., 2018; Wolke et al., 2009). The latter is particularly challenging for research of pregnancy exposures on the development of mental health problems, as it requires participation until children are old enough to properly assess mental health phenotypes. Furthermore, research has found that smoking, lower education, and other factors related to lower SEP increase chances of dropout from longitudinal studies (Taylor, Jones, et al., 2018; Wolke et al., 2009). Thus, associations found in observational research of smoking and caffeine consumption during pregnancy and offspring mental health outcomes are likely biased due to selection bias, the direction of which is difficult to predict. Statistical methods such as multiple imputations and controlling for known confounding variables can be useful approaches to account for selection bias. The former approach assumes that the pattern of missing data is missing at random, which is rarely the case in studies investigating mental health phenotypes (Hammerton & Munafò, 2021). Statistically adjusting for confounding variables in complete case analyses may yield unbiased results even if the data is not missing at random, given that the adjustment resulted in the outcome being independent of the likelihood of being a complete case (Hughes et al., 2019).

1.8.2 Measurement bias

Measurement bias occurs when measures of the behaviour of interest are imprecise and capture additional variance of other variables that can dilute the effect that we are interested in (Hammerton & Munafò, 2021). This type of bias is commonly present in self-report measures that assess socially stigmatised behaviours (Hammerton & Munafò, 2021). Prevalence estimates of smoking during pregnancy are likely to be an underestimation of the actual percentage of mothers who smoke during pregnancy. Particularly during pregnancy, self-reported smoking is likely to suffer from response bias with mothers underreporting smoking because of fear of social stigmatisation. This has been demonstrated by studies that found self-reported smoking during pregnancy to be underreported 20% to 25% of the time when compared to urinary cotinine levels

as a biomarker for actual smoking (England et al., 2007; Shipton et al., 2009). Even though caffeine consumption during pregnancy is less stigmatised than smoking, there is evidence for some social stigmatisation of caffeine consumption during pregnancy that may cause mothers to underreport actual consumption (see section 1.9.2 on mother blaming). Furthermore, caffeine consumption in research studies is typically imprecisely measured with most studies asking for self-report of cups of caffeinated beverage without considering variable cup sizes (van Dam et al., 2020). Particularly during pregnancy, most research has focussed on caffeine consumption from coffee only without differentiating between different types of coffee (e.g., espresso vs. filter coffee, etc.) and neglecting other sources of caffeine (caffeinated soda drinks), as well as effects of other ingredients than caffeine that are contained in coffee (CARE Study Group, 2008; Chen et al., 2014; Grosso & Bracken, 2005).

1.8.3 Confounding bias

One of the main problems for deriving causal inferences from observational research is confounding. The confounding structure of smoking and caffeine consumption during pregnancy is highly complex, making it extremely complicated to narrow down whether an observational association with offspring mental health problems is truly due to the intrauterine smoking or caffeine exposure. The next sections will summarise the main confounding factors of observational associations between maternal smoking and caffeine consumption during pregnancy and offspring mental health problems.

1.8.3.1 Social pattern of smoking and caffeine consumption during pregnancy

Smoking and high caffeine consumption during pregnancy are strongly correlated and tend to also correlate with other substance use behaviours such as alcohol consumption (Chen et al., 2014; Loomans et al., 2012; Robinson et al., 2010), which is known to cause developmental problems, including symptoms of mental health problems, under the umbrella of Fetal Alcohol Spectrum Disorder (Lange et al., 2017). Despite the correlation between smoking and caffeine consumption during pregnancy, studies report different social patterns of the individual behaviours. Generally, studies report that mothers who smoke during pregnancy tend to be younger and to have a lower SEP than mothers who do not smoke

during pregnancy (Chamberlain et al., 2017; Schneider & Schütz, 2008; Smedberg et al., 2014). In contrast, studies that investigated caffeine consumption during pregnancy report that mothers who consume caffeine during pregnancy are older and have a higher SEP than women who consume less or no caffeine during pregnancy (Berglundh et al., 2020a; Loomans et al., 2012; Weng et al., 2008). However, a recent study reported an inverse association between caffeine consumption during pregnancy and SEP (income and education) (Patti et al., 2021). As noted earlier, there are only few nationally representative studies with data on maternal caffeine consumption during pregnancy and the cohort studies mentioned above may not be fully representative of caffeine consumption during pregnancy in the general population (Verster & Koenig, 2018). Furthermore, there is some indication that consumption of different caffeinated drinks is associated with different socio-demographic characteristics. For instance, whereas caffeinated coffee consumption during pregnancy was associated with older maternal age (> 35 years), consumption of caffeinated tea and soda consumption during pregnancy was associated with younger maternal age (< 20 years) (Chen et al., 2014).

As mentioned above, the majority of women drink less than 200 mg of caffeine per day and it is likely that mothers, who consume more than 200-300 mg a day of caffeine, are demographically more similar to mothers who smoke during pregnancy, than mothers who consume lower amounts of caffeine during pregnancy (Chen et al., 2014). There is some indication that mothers who consume high levels of caffeine, on the one hand tend to be older and higher educated, on the other hand tend to work and smoke more, and drink more alcohol during pregnancy, compared to mothers who consume lower amounts of caffeine during pregnancy (Bech et al., 2007; Chen et al., 2014; Loomans et al., 2012).

For smoking, the negative health effects on offspring are generally known and pregnancy guidelines across countries strictly recommend abstaining completely during pregnancy. Continuing to smoke during pregnancy is associated with social stigmatisation and may be correlated with other variables that may contribute to the risk for mental health problems in offspring. Across studies, it has been found that mothers who continued smoking during pregnancy face

severe hardships during pregnancy compared to mothers who do not smoke or manage to quit during pregnancy (DiClemente et al., 2000; Massey & Compton, 2013; Robinson et al., 2010; Roza et al., 2009; Schneider & Schütz, 2008; Smedberg et al., 2014). In comparison, they tend to have lower SEP, be younger, live on their own or be in a relationship with someone who smokes, have more children, receive insufficient prenatal care, experience more stress, physical violence and unemployment, and to have more mental health problems (DiClemente et al., 2000; Massey & Compton, 2013; Robinson et al., 2010; Roza et al., 2009; Schneider & Schütz, 2008; Smedberg et al., 2014). Further, in comparison to mothers who continued smoking during pregnancy, regular smokers, who managed to quit during pregnancy, were found to have offspring with an easier temperament and to have more perceived self-efficacy three years post-pregnancy than mothers who did not manage to quit (Robinson et al., 2010).

1.8.3.2 Maternal mental health problems and smoking and caffeine consumption during pregnancy

An important domain that may confound associations between smoking and caffeine consumption during pregnancy and offspring mental health problems is maternal mental health. In non-pregnant populations, both, tobacco and caffeine consumption are associated with mental health problems (Alasmari, 2020; Kendler et al., 2008; Lara, 2010). In contrast to caffeine consumption during pregnancy, the relationship between smoking during pregnancy and maternal mental health problems has been thoroughly reported. A national survey conducted in the US found that smoking amongst pregnant women (N = 5,442) was associated with experiencing more severe psychological distress across all socio-demographic groups (Goodwin et al., 2017). Women who have experienced severe psychological distress within the past month were three times more likely to smoke during pregnancy compared to women who did not experience acute psychological distress, even after adjusting for socio-demographic variables. The risk for smoking was lower for women who experienced severe psychological distress at some point over the last year but not within the past month; yet still twice as large compared to women who did not experience any severe psychological distress, even after controlling for socio-demographic variables. Importantly, severe psychological distress tremendously increased the prevalence

of smoking during pregnancy for women with a higher socio-demographic background (high education and income), who generally are found to be less likely to smoke during pregnancy. Whereas less than 2% of women with a college degree and without having experienced severe psychological distress reported smoking, 45% of women who have a college degree and have experienced psychological distress reported smoking during pregnancy. However, no relationship between severe psychological distress and smoking heaviness during pregnancy was observed. These results are in line with studies from the Netherlands (N = 1,947) (Goedhart et al., 2009), Norway (N = 90,040) (Moylan et al., 2015; Zhu & Valbø, 2002), and Scotland (N = 395) (Pritchard, 1994), which reported associations between maternal smoking during pregnancy and a higher risk for depression and anxiety symptoms. Two common theories for the high comorbidity between mental health problems and substances use are: (1) self-medication, where people already suffering from mental health problems use caffeine and tobacco to cope with their symptoms (Khantzian, 1997; Morgan & Jorm, 2008), and (2) the actual smoking and caffeine consumption behaviour contributing to the development of mental health disorders (Wootton et al., 2018). The associations between smoking during pregnancy and maternal mental health problems raise the question whether associations observed between prenatal exposure to smoking and offspring mental health problems may be (partly) explained by a shared genetic predisposition to smoking and mental health problems or because of shared environmental risk factors between smoking and mental health problems. Regardless of the true direction of effect, smoking is commonly perceived as a stress relief amongst smokers and abstaining from smoking and/or cutting down caffeine during pregnancy might increase stress (e.g., by experiencing unpleasant withdrawal symptoms). Considering the differences between socioeconomic variables of smokers and non-smokers outlined above, a pregnancy might pose as an additional strain to already challenging circumstances, such as financial insecurities and relationship problems, making a behaviour change even more challenging (DiClemente et al., 2000). This is in line with research that found women with mental health problems who are pregnant and smoke, despite high acceptance of referral to smoking cessation programs, to be less likely to actually stop smoking during pregnancy than pregnant women without mental health problems (Howard et al.,

2013). Research on the association between maternal caffeine consumption and mental health problems during pregnancy is scarce and even studies investigating the effect of intrauterine caffeine exposure on offspring mental health problems rarely report associations between maternal mental health during pregnancy (Bekkhus et al., 2010; Berglundh et al., 2020a; Klebanoff & Keim, 2015; Mikkelsen et al., 2017). Those studies that did report association between maternal caffeine consumption during pregnancy and maternal mental health found a positive association with increased worry (Mourady et al., 2017), depression and nervousness (Del-Ponte et al., 2016). In contrast, one study reported a negative association between maternal caffeine consumption and maternal risk for psychopathology (Linnet et al., 2009). Adding to the risk of confounding through exposure to maternal mental health problems during pregnancy, postnatal exposure may be an additional risk for confounding. Maternal mental health problems during pregnancy are strongly correlated with maternal and paternal postnatal mental health problems (Beck, 2001; O'hara & Swain, 1996), which are positively associated with the risk for offspring mental health problems (Aktar et al., 2019).

1.8.3.3 Offspring's own exposure to smoking or caffeine

As outlined above, there are substantial differences between mothers who smoke and/or consume caffeine during pregnancy and mothers who do not. The list of confounding variable becomes even longer and more complex, with increasing the temporal gap between the actual intrauterine exposure and the assessment of the mental health problem (Davey Smith, 2008). For instance, studies that investigate the effect of maternal smoking on offspring mental health problems in adolescence could be confounded through the effects of offspring's own smoking or caffeine consumption on their mental health (either due to a shared genetic liability or offspring modelling maternal smoking behaviour, or both). Therefore, in this thesis I focus on identifying early manifestations of mental health problems in young children, where own smoking or caffeine consumption is unlikely to be confounding associations with maternal smoking and caffeine consumption during pregnancy.

1.9 Why we need better evidence

Despite issues with confounding and other types of bias, observational evidence currently drives our understanding of the role of prenatal exposures and the risk for offspring mental health outcomes. This weak evidence base is used to support public health policy and advice to pregnant women. Therefore, there is an urgent need to improve the evidence base by distinguishing causal relationships from mere correlation and by contextualising the role of maternal pregnancy exposures against other factors that could affect mental health, such as paternal exposures.

1.9.1 Pregnancy guidelines

There is a wealth of advice and guidance for health behaviours during pregnancy to ensure the best developmental outcome for the unborn child. Current UK guidelines advise women to quit smoking immediately once they find out about their pregnancy, as well as stressing that every additional cigarette poses a threat to their baby's health (https://www.nhs.uk/pregnancy/keeping-well/stop-smoking/). This guideline is based on scientific evidence that found smoking during pregnancy to be associated with negative birth outcomes such as premature birth, stillbirth, low birth weight, and sudden infant death syndrome (Marufu et al., 2015; Salihu & Wilson, 2007; Tyrrell et al., 2012). Furthermore, the guideline is based on research indicating an increased risk for asthma and overall health problems later in life (Hylkema & Blacquiere, 2009). For other maternal health behaviours during pregnancy, the evidence of harm to the fetus is much weaker. Generally, guidelines on these behaviours follow the "precautionary principle" and thus even if clear empirical evidence of the effect of an exposure during pregnancy on offspring is lacking, it is still recommend abstaining from the behaviour.

In contrast to smoking during pregnancy, there are no uniform guidelines for caffeine consumption during pregnancy yet, as they differ from country to country. A study that systematically screened guidelines for caffeine consumption as published on the website of the "Food and Agriculture Organization" of the United Nations (http://www.fao.org/nutrition/nutrition-education/food-dietary-guidelines/en/) reports that out of 90 countries, 81 note caffeine consumption in their dietary guidelines, of which the majority (80%) were published over the past

10 years (Reyes & Cornelis, 2018). Yet only 13 countries provided advice for pregnant women. All of the pregnancy guidelines recommend avoiding caffeine use during pregnancy and eight countries further specified to limit caffeine consumption to no more than 200-300 mg/day. These pregnancy guidelines of limiting caffeine consumption are in line with the recommendation of a panel of the EFSA that, after reviewing the scientific evidence, has declared up to 200 mg/day of caffeine to be safe for consumption during pregnancy (EFSA Panel on Dietetic Products, Nutrition & Allergies (NDA), 2015). The conclusion was derived based on results of prospective birth cohorts that reported a dose-response relationship with caffeine consumption during pregnancy and risk for adverse birth outcomes related to low birth weight (e.g., small for gestational age) (Bakker et al., 2010; CARE Study Group, 2008; Sengpiel et al., 2013).

There are very good reasons for following the precautionary principle when there is uncertainty about the effect of an exposure during pregnancy – abstinence is likely to be the safest option for mother and fetus. A major challenge for developing pregnancy guidelines is the need to balance the lack of causal evidence for behaviours during pregnancy with harm avoidance for offspring and the invasiveness of the intervention for maternal autonomy (Stratil et al., 2020). The degree of precaution within pregnancy guidelines should be transparently communicated and the research that alerts for precaution should be appropriately referred to and summarised. For instance, the UK guideline for alcohol consumption during pregnancy highlights the level of uncertainty about safe amounts for alcohol consumption during pregnancy:

"*Experts are still unsure exactly how much – if any – alcohol is completely safe for you to have while you're pregnant, so the safest approach is not to drink at all while you're expecting.... Women who find out they're pregnant after already having drunk in early pregnancy should avoid further drinking. However, they should not worry unnecessarily, as the risks of their baby being affected are likely to be low*"

(https://www.nhs.uk/pregnancy/keeping-well/drinking-alcohol-while-pregnant/)

The precautionary principle may appear as the easiest strategy to mitigate the risk for offspring's health problems; however, it may also decrease the urgency to

establish guidelines based on causal evidence (Gignon et al., 2013). It is commonly perceived that abstaining from smoking or caffeine consumption during pregnancy is only a small cost in comparison to the potential risk of compromising offspring's health and its associated burden for society. However, despite a substantial amount of research indicating that women who smoke during pregnancy face sever hardships and lack a social support system there seems to be a lack of understanding for the wider social-structural context these behaviours are embedded in and lack of compassion in society for mothers who struggle with abstaining. Pregnancy guidelines that do not clearly state and communicate the level of uncertainty, may be misinterpreted by the public as if there was substantial evidence and increase stigmatisation of women not adhering to pregnancy guidelines. In sum, precautionary principle can limit women's autonomy, especially regarding the stigmatisation of behaviours that are socially and culturally engrained and widely accepted in non-pregnant adults.

1.9.2 Mother blaming

The stigmatisation of certain behaviours in pregnant women can contribute to a culture of "mother blaming", which includes the phenomenon of society (including health professionals and mothers themselves) holding the mother accountable for any (psycho-) pathology of their children, as well as for her own life circumstances, such as poverty (Koniak-Griffin et al., 2006). An interesting study from Australia compared the perception of smoking and non-smoking pregnant women amongst college students. The study found that, whereas women were overall perceived more positively (healthier, more accepting, proud, empowered, believing, and relaxed) if they were pregnant compared to non-pregnant, this perception changed towards the opposite when given the information that the women was smoking. Pregnant smoking women were perceived as more ignorant and selfish than non-pregnant smokers (Wigginton & Lee, 2013). This highlights the social stigmatization of women who smoke during pregnancy, which can be problematic because it prevents mothers to honestly report their smoking behaviour during pregnancy (see section 1.8.2) and to seek assistance to stop smoking (Koniak-Griffin et al., 2006; Wigginton & Lee, 2013).

Compared to smoking, there is less social stigma around caffeine drinking in pregnancy, but changes in public health guidelines could change that. For instance, a recently published narrative review has claimed that there would be no safe level of caffeine for women to consume during pregnancy and that a substantial number of miscarriages in the US might be attributable to maternal caffeine consumption during pregnancy (James, 2020). This narrative review quickly attracted media attention and resulted in polarizing headlines, which likely have caused confusion, anxiety and guilt in many women as well as may have contributed to an overestimation of the known risk of caffeine during pregnancy among the general public (e.g., *No Safe Level of Coffee Drinking for Pregnant Women, Study Says*, 2020). My supervisors and I contributed to the reply by Murphy and colleagues to the editor of the narrative review of James (James, 2020), in which we highlighted the limitations of the research and alerted about jumping to conclusions based on observational evidence (Murphy et al., 2020). Our main criticisms about the scientific conduct of the narrative review were: (1) the non-systematic search for literature, (2) the restricted discussion of the limitations of the studies included, (3) the premature causal inferences made from observational research and superficial discussion of potential confounding, and (4) the lack of information about the effect size of caffeine during pregnancy on offspring outcomes. In addition, concerns were raised about the incomprehensive press release of the paper that led to extrapolated media headlines and the false representation of pregnant women recklessly consuming caffeine during pregnancy without taking into account the wider social structures. Professor James replied to our letter, rejecting the scientific arguments that support concerns of bias in his research. Instead, he criticised the expertise of Murphy and the other 20 co-authors, despite all being experts in relevant areas of caffeine consumption during pregnancy (researchers in epidemiology, pregnancy exposures, psychology, and health practitioners in the field of prenatal care, etc.), and suggested bias because of high chances of Murphy and her co-authors being caffeine consumers themselves:

"It is reasonable to assume that most, if not all, of Murphy and her 20 co-authors are habitual caffeine consumers, sharing common features with other consumers, including physical dependence and desire to continue to consume. The self-

serving bias therein is likely to encourage resistance to information that challenges habit".

Furthermore, James did not address the concerns about the inadequate press release that led to polarized media headlines and the associated harm that such headlines may have caused for pregnant women. In sum, the article of James (James, 2020) and the surrounding scientific debate reflects the confusion that arises, even within the scientific community, when only observational research evidence is available, and highlights the confusion and scepticism that the communication of such research to the general public and expectant parents may bring about.

Polarizing media articles and overblown interpretations of the scientific evidence neglect the psychological wellbeing of the mother as well as an understanding of the wider social and structural determinants of health. In particular, the contribution of a partner's health behaviours pre-conception and their influence on maternal wellbeing and health during pregnancy is often ignored or downplayed (Román-Gálvez et al., 2018). There is emerging evidence that partner smoking is a strong predictor of continued maternal smoking during pregnancy and that partners are unlikely to stop smoking during pregnancy (Riaz et al., 2018; Román-Gálvez et al., 2018; Schneider et al., 2010). Overall, having more evidence-based pregnancy guidelines and a less toxic risk communication for behaviours during pregnancy would help to reduce anxiety and increase well-being for parents and may therefore contribute towards a healthy development for offspring. The analyses of this thesis are aimed at contributing towards more evidence-based pregnancy guidelines by contextualising the observational evidence base with evidence from causal inference techniques using genetic and epigenetic data.

1.9.3 Testing, rather than reinforcing, the DOHaD hypothesis

To establish evidence-based pregnancy guidelines, it is important that researchers, who investigate intrauterine effects on offspring health, become aware of possible societal biases that may influence the hypothesis generation, conduction, and communication of their research. Sharp and colleagues (Sharp, Lawlor, et al., 2018) argue that the damaging culture of "mother blaming" is in part explained by deeply held assumptions about the "causal primacy" of maternal pregnancy

effects – the pervasive idea that maternal exposures during pregnancy are the most important factors determining child health, and that other factors (e.g., paternal factors, postnatal exposures and wider societal influences) are much less important. Supporting this, in a 2019 article, we illustrated that the traditional focus of DOHaD research has been predominantly on maternal exposures during pregnancy (Sharp et al., 2019). For the article I screened and extracted the relevant information of all the original research articles that were published in the "Journal of the Developmental Origins of Health and Disease", up until January 2019. Comparing the number of articles that focussed on maternal exposures with articles that focussed on paternal exposures, for both, animal (N studies = 144) and human studies (N studies = 182) we found that results were clearly biased towards the focus on maternal pregnancy exposures. The sunbursts displayed in the paper illustrate our findings (Figure 1.4, interactive version of the sunburst can be found on https://gs8094.shinyapps.io/sunburst/).

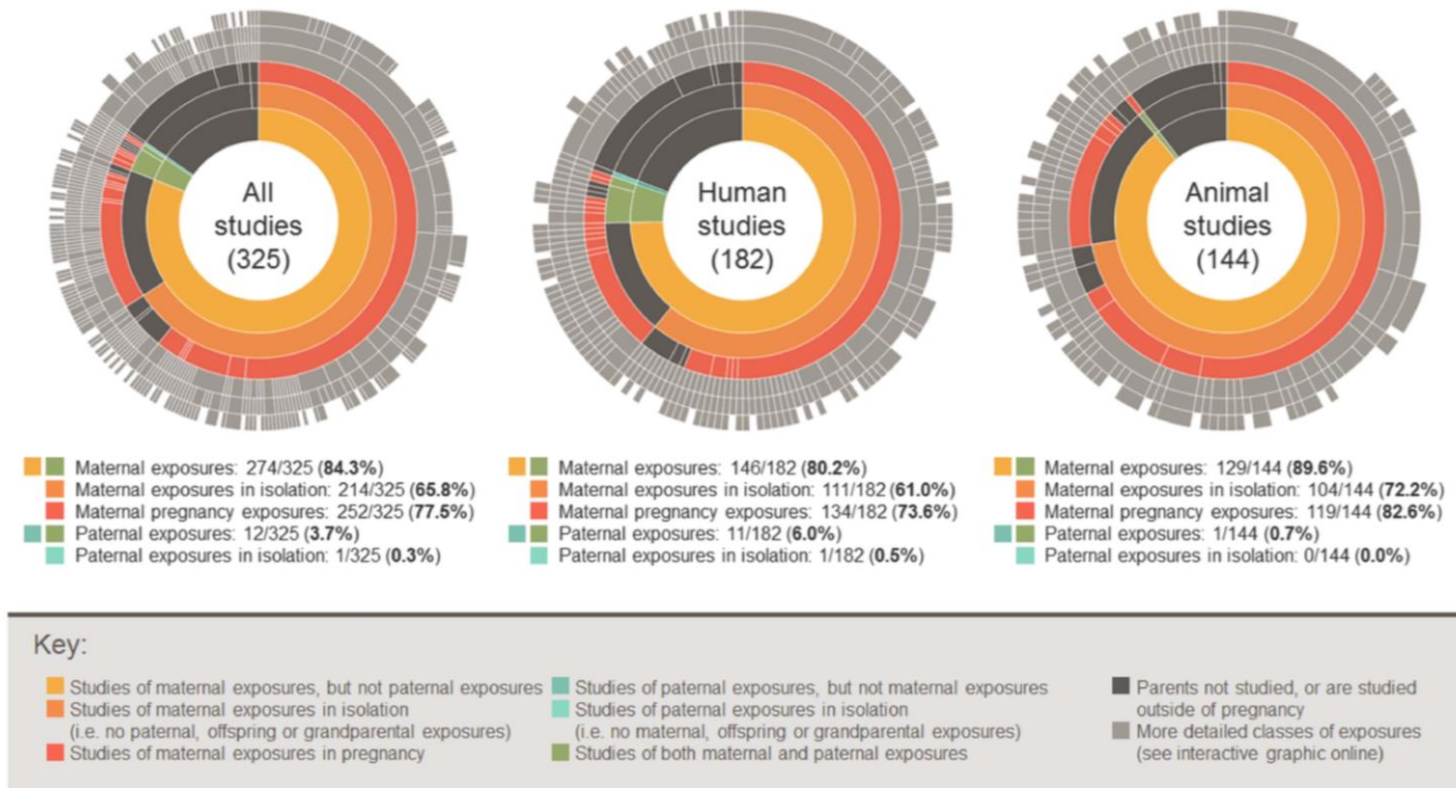


Figure 1.4. Sunburst of studies published in the Journal of the Developmental Origins of Health and Disease and the proportions of their focus on maternal and paternal exposures *(credit: Sharp et al., 2019).*

As shown in Figure 1.4 in the left sunburst, across human and animal studies (N studies = 325), 84% of the studies investigated maternal exposures (at any time, not just pregnancy) and 78% focussed on maternal exposures during pregnancy. Yet only 4% of studies investigated paternal exposures (at any time, not just pregnancy). Further, whereas 81% of studies focussed exclusively on maternal exposures (removing studies that included paternal, offspring, grandparental exposure, etc.), only 0.3% (one study) exclusively focussed on paternal exposure (removing studies that included maternal, offspring, grandparental exposure, etc.).

On the surface, there are good reasons for this, including the unique biological relationship between a pregnant mother and fetus, and the relative practical ease of recruiting pregnant women to birth cohorts during antenatal appointments. Against the latter argument stands the observation that in the DOHaD literature there was no difference in bias towards maternal pregnancy exposures when restricting studies to animal data only (Figure 1.4, right sunburst; 90% studying maternal exposures vs. 0.7% studying paternal exposures). This illustrates that lack of paternal data is not driving the bias observed in human studies (Figure 1.4, middle sunburst; 80% studying maternal exposures vs. 6% studying paternal exposures). This imbalance in the DOHaD literature supports the public focus on the behaviours of pregnant mothers and can be used to support further research to study more maternal pregnancy effects, acting via a looping effect to further intensify the focus (Sharp, Lawlor, et al., 2018). The DOHaD hypothesis is therefore further reinforced rather than tested. When this is considered alongside the fact that much of the evidence in DOHaD is weak and correlational, the current overfocus on maternal pregnancy exposures is less justifiable. There is now a need to redress the imbalance in DOHaD research by using causal inference approaches and contextualising the evidence by studying other factors (like paternal factors) with a similar frequency. This PhD project aims at complementing the current observational evidence base for effects of smoking and caffeine consumption during pregnancy on offspring mental health outcomes by applying genetic and epigenetic causal inference methods. The next section reviews available causal inference techniques that can be applied to answer research questions for which RCTs are difficult to implement.

1.10 Causal Inference Techniques

The scarcity of controlled human trials in the field of prenatal exposures makes it difficult to isolate the effect of a single exposure on an outcome (Knopik et al., 2018). As shown above, mothers who smoke and consume caffeine during pregnancy are facing many psychosocial adversities, which likely also influence offspring's psychological development. Statistically adjusting for known confounding variables in observational research is unlikely to be sufficient to account for the highly complex underlying confounding structures (Robinson et al., 2010). Studies have shown that adjustment for known confounding factors of prenatal smoking and mental health problems still leads to biased results and genetically sensitive designs are needed to account for inherited confounders within families (Thapar & Rutter, 2009). Causal inference techniques using genetic and paternal data are thus far the best methods to account for a large proportion of confounding variables, including unknown and unmeasured confounding variables.

1.10.1 Negative control studies

Studies that use paternal exposures measured during their partner's pregnancy as negative control exposures can help to isolate intrauterine effects by providing indication for the presence of confounding in associations between maternal behaviours during pregnancy and offspring health outcomes (Taylor, Davey Smith, et al., 2014). This approach assumes that maternal and paternal behaviours have similar confounding structures, but that there is little or no biological plausibility for a causal effect of the paternal exposure. For instance, maternal and paternal smoking during pregnancy have been found to share a similar socioeconomic confounding structure (Roza et al., 2009). If in addition to maternal data, paternal data on smoking and caffeine consumption is available, the comparison of the magnitude of the maternal-offspring and paternal-offspring effect estimates allows making inferences about a potential intrauterine effect (Gage et al., 2016). If a direct intrauterine effect exists, one would expect the maternal-offspring effect estimate to be larger than the paternal-offspring effect estimate, as a direct biological effect of smoking and caffeine consumption during pregnancy is only plausible through the maternal but not paternal environment.

Even if the mother's partner smokes during pregnancy, the effect estimate of maternal smoking during pregnancy on offspring outcomes would still be expected to be larger than the effect through partner second-hand smoking during pregnancy (Taylor, Davey Smith, et al., 2014). Therefore, if paternal data during pregnancy is available, negative control studies can be used to investigate whether the observed association between maternal exposure and offspring outcome is confounded or reflecting a true intrauterine effect. The imbalance in research focussing on maternal compared to paternal pregnancy exposures, that was discussed in section 1.9.3, prevents investigation of intrauterine effects using negative control designs (Easey & Sharp, 2021).

Some studies were able to apply a negative control design to investigate intrauterine effects of smoking during pregnancy on offspring mental health outcomes. Yet, comparison between maternal versus paternal associations of smoking during pregnancy and offspring externalising problems show inconclusive results about a true intrauterine effect (Dolan et al., 2016). Whereas some negative control studies found a stronger effect between maternal smoking and offspring externalising problems than for paternal smoking during pregnancy (Brion et al., 2010; Nomura et al., 2010), others did not (Langley et al., 2012; Roza et al., 2009). For internalising problems, a meta-analysis of negative control studies of five independent birth cohorts (total N = 21,246) has not found evidence for an intrauterine effect of smoking during pregnancy on offspring depressive symptoms in adulthood (Taylor et al., 2017). Despite offspring of mothers who smoked during pregnancy having higher odds for adult depression, the odds ratio was not statistically different from the odds ratio of partner smoking during pregnancy and offspring depressive symptoms. Within the same study, a sibling-comparison analysis (N = 258) was conducted that did not find a difference in risk for depression between siblings that were exposed to smoking in utero and their sibling who had not been exposed. Together, results of the negative control meta-analysis and sibling comparison study suggest that maternal smoking during pregnancy is not causally associated with an increased risk for offspring depression and that previous reported associations might be explained by confounding factors.

1.10.2 Mendelian Randomisation (MR)

Another approach to overcome some of the limitations of observational research, such as unmeasured confounding and reverse causation, is to use genetic variants as instrumental variables that can be used as a proxy for the behaviour of interest. This method is called Mendelian Randomisation (MR) because it draws on Mendel's law of random segregation, which posits that alleles are randomly segregated and assorted at conception. Genetic variants that proxy for the behaviour of interest should therefore not be related to common confounding variables, and thus can be used to mimic a randomised controlled trial in observational research (Davey Smith & Ebrahim, 2003a). MR yields less biased estimates under the premise that the instrumental variable assumptions hold which require that (1) the genetic variants are strongly associated with the exposure of interest, (2) the genetic variants are not associated with any variables affecting the outcome, and (3) the genetic variants are only associated with the outcome through the exposure of interest (Taylor, Davies, et al., 2014).

1.10.2.1 MR studies of smoking and caffeine consumption and mental health problems

In the general population, MR studies have helped to shed light on the contribution of smoking to the risk for mental health problems. For instance, MR studies indicate that smoking increases the risk to develop bipolar depression (Vermeulen et al., 2019) and schizophrenia (Wootton et al., 2018). In contrast, MR studies of ADHD and depression support the self-medication hypothesis, with people with ADHD and depression symptoms being more likely to initiate smoking but little evidence for smoking increasing the risk for ADHD or depression (Lewis et al., 2011; Treur et al., 2019). MR can also be applied in the context of prenatal exposures on offspring outcomes, a method called prenatal MR (Diemer et al., 2020), or intergenerational MR (Lawlor et al., 2017). A recent systematic review has identified 43 studies that have applied such a design, by using maternal genetic variants to proxy for the exposure of interest, and tested their associations with offspring outcomes (Diemer et al., 2020). The review identified four studies that have investigated the effects of prenatal smoking, of which the only mental health outcome investigated was autism spectrum disorder that showed weak evidence for an effect (Caramaschi et al., 2018; Diemer et al.,

2020). Of the 43 studies, only one study has investigated the effects of maternal caffeine consumption during pregnancy on stillbirth. Also, in non-pregnant populations, only few studies have used MR to investigate potential causal effects of caffeine on mental health outcomes (Cornelis & Munafo, 2018). One study has investigated the effect of caffeine on depression and found no evidence for an effect (Kwok et al., 2016). Another study testing associations between the genetic variants for coffee and all ICD recorded health problems in the UK biobank, including 43 mental diseases, also did not find support for an association between coffee and mental health outcomes (Nicolopoulos et al., 2020). In this thesis, I use methods based on the principles of MR (Chapter 3: PheWAS; Chapter 6: 2-sample MR) in order to make inferences about whether there is an intrauterine effect of smoking and caffeine exposure on mental health outcomes.

1.10.2.2 MR studies of the relationship of smoking and caffeine consumption

Few studies have investigated the relationship of smoking and caffeine consumption (Cornelis & Munafo, 2018). Results of these studies suggests that smoking might have a causal effect on increasing caffeine consumption (Bjørngaard, Nordestgaard, Taylor, Treur, Gabrielsen, Munafò, Nordestgaard, Åsvold, Romundstad, & Smith, 2017) but little evidence for caffeine consumption increasing smoking (Ware et al., 2017). Yet, one study did not find evidence for an association between smoking and caffeine consumption (Verweij et al., 2018). In sum, there is some evidence that smoking is causally increasing caffeine consumption and thus there might be an interactive effect of smoking and caffeine exposure during pregnancy on offspring mental health. Before such an interaction effect is further investigated, it is important to increase understanding of the association between prenatal caffeine exposure and mental health outcomes. Due to the scarcity of research on the effect of intrauterine caffeine exposure on offspring mental health, this thesis will focus on the individual effects of smoking and caffeine consumption during pregnancy and not on a potential interactive effect of smoking and caffeine.

1.10.2.3 Instruments for MR studies of smoking and caffeine consumption

The availability of large genome-wide association studies (GWAS) has made it possible to discover thousands of common genetic variants (single nucleotide polymorphisms; SNPs) that can be used as instruments in MR. GWAS are

considered hypothesis generating, as in contrast to candidate-gene studies, the entire genome is tested for associations with a phenotype (Manolio et al., 2009). The GWAS method has led to the discovery of many novel genetic variants that are associated with complex traits and illuminated potential biological pathways to mental health problems (Visscher et al., 2017). Further, they contributed towards establishing genetic instruments for modifiable risk exposures, such as smoking and caffeine consumption, that can be applied in MR studies (Liu, Jiang, Wedow, Li, Vrieze, et al., 2019; The Coffee and Caffeine Genetics Consortium et al., 2015). Genetic variants associated with smoking and caffeine consumption have shown to predict 1% to 4% of variance of these behaviours in the general population but their use for predicting these behaviours during pregnancy is still unknown (Lawlor et al., 2017). Due to the high number of tests, which results from testing each SNP for association with a phenotype, GWAS require large sample sizes to detect effects and therefore often include meta-analysis of different prospective cohort studies. However, considering the many novel variants that have been identified in GWAS, little is understood about the molecular mechanisms that explains these gene-phenotype associations (van der Sijde et al., 2014). In this thesis, I use genetic instruments derived from GWAS of smoking and caffeine consumption to explore whether observational associations between maternal smoking and caffeine consumption and offspring outcomes (DNA methylation and mental health) are confounded, or indeed indicating a causal effect of the exposures.

1.10.3 Other genetically sensitive designs to infer causality

Other examples of study designs that adjust for shared genetic factors between mother and offspring are studies with data on Assisted Reproductive Technologies and children-of-twins (Thapar & Rutter, 2009). Thapar and colleagues (Thapar et al., 2009) compared ADHD symptoms of offspring who have been exposed to smoking during pregnancy by their genetically related or unrelated mother (oocyte/embryo donations, surrogate mothers). The results showed that the magnitude of intrauterine smoke exposure on offspring ADHD was larger for related than unrelated mother-child pairs, indicating that inherited common genetic factors are more likely to be responsible for offspring ADHD symptoms than the actual intrauterine smoke exposure. However, while these

study designs are highly powerful, as they allow us to account for and differentiate between unmeasured environmental and genetic confounders, they require data that is scarce and difficult to obtain, in contrast to negative control and MR studies that were applied in this thesis.

1.10.4 The meet-in-the-middle approach to establish causality

The analyses of this thesis will follow the *meet-in-the-middle* approach (Vineis et al., 2013) to establish whether there is a causal effect of smoking and caffeine consumption during pregnancy on offspring mental health outcomes. According to the meet-in-the-middle approach, omics data provide an opportunity to derive causal pathways by discovering intermediate biomarkers that link an exposure to the risk of disease (Richmond et al., 2014; Vineis et al., 2013; Vineis & Perera, 2007). Data of prospective birth cohorts can be used to discover biomarkers for the exposure in a prospective fashion and retrospectively investigate whether these biomarkers are enriched in people who developed a certain disease compared to people who did not develop the disease. Vineis and colleagues (Vineis et al., 2013) summarise the meet-in-the-middle approach in three steps: The first step examines the relationship between the exposure and disease. Next, the association between the exposure and a potential biomarker for the exposure (in this case, DNA methylation) is investigated. In the last step, the relationship between the biomarker and the disease is tested (Vineis et al., 2013). If the evidence for associations is congruent in the analyses of each step, confidence in a causal effect between the exposure and outcome increases.

The three steps are used to investigate a potential causal effect of prenatal smoking and caffeine exposure in this thesis by: First, exploring the association between maternal smoking and caffeine consumption on offspring mental health outcomes in the phenome-wide association study (Chapter 3). Second, investigating associations between prenatal caffeine exposure and offspring cord blood DNA methylation (Chapter 4). Third, the association between prenatal smoking and caffeine associated DNA methylation changes and the risk for offspring internalising problems are examined through an EWAS meta-analysis of internalising problems (Chapter 5), an enrichment analysis and a two-sample MR analysis (Chapter 6). A causal effect of prenatal smoking and caffeine exposure on offspring mental health would be supported by results showing evidence for

associations between: (1) maternal smoking and caffeine consumption on offspring mental health (exposure-disease association), (2) maternal prenatal smoking and caffeine consumption on offspring DNA methylation (exposure-intermediate biomarker association), and (3) prenatal smoking and caffeine associated DNA methylations changes and offspring internalising problems (intermediate biomarker-disease association).

1.10.5 Triangulation

As shown by the studies reported above, none of the different methods can yield a clear answer to the question whether maternal smoking and caffeine consumption during pregnancy increase the risk for offspring mental health problems. The concept of triangulation acknowledges that no single research method or design can provide the answer to a causal question. Rather it highlights the value in exploring the same research question through different techniques, subject to different or opposing biases, in order to understand causal effects (Lawlor et al., 2016). In contrast to meta-analyses and replication studies, which aim at synthesising research approaches that are ideally very similar to one another, triangulation takes advantage of the diversity of underlying biases. If results of different research methods and designs, that bias results in different directions, are congruent in their conclusions about effects of an exposure on an outcome, confidence in causality increases (Pingault et al., 2018). The criteria for triangulating evidence for a causal effect are that: (1) at least two different methodologies are compared, which are subject to different and vastly independent sources of bias, (2) there is a common underlying research question, and timing and duration of the exposure is considered, and (3) the main source of bias of each method is understood and recognized, and the direction of bias is taken into account (ideally the different methods bias results in opposing directions). As outlined earlier, the main sources of bias in observational research are selection, measurement, and confounding biases (Hammerton & Munafò, 2021; Lawlor et al., 2004). By contextualising observational research evidence about smoking and caffeine consumption during pregnancy on offspring mental health outcomes, which is likely confounded by socioeconomic factors, with MR studies that are free of confounding (but show different biases), it could be distilled whether a true causal relationship exists. In addition to applying

statistical approaches and using particular study designs, the identification of a potential biological pathway that could explain how prenatal exposure to smoking and caffeine consumption could become biologically embedded to increase the risk for offspring mental health problems would also strengthen the causal evidence base (Vineis et al., 2013).

1.11 Chapter summary and conclusion

In this thesis, I am attempting to improve the causal evidence base regarding the effects of maternal smoking and caffeine consumption during pregnancy on offspring mental health outcomes. Over the last decade, awareness for the importance of protecting mental health, particularly amongst young people, has tremendously increased. Data from the UK has found that approximately one out of eight children, between the age of 5 and 19, fulfil criteria for a mental health diagnosis (Sadler et al., 2018). This warrants a better understanding of the factors that influence mental health problems in children to improve detection of early indicators for mental health problems. Most DOHaD research has focussed on investigating maternal behaviours during pregnancy as a risk factor for offspring health problems but causal evidence is scarce (Sharp et al., 2019).

In this chapter, I have outlined how the high prevalence of smoking and caffeine consumption during pregnancy warrants a more detailed understanding of the range and mechanisms of effects on offspring health outcomes. I have highlighted several imbalances and gaps that are currently present in the field and that I attempt to address within this thesis. First, most current evidence on associations between maternal health behaviours and child mental health pertains to maternal smoking and there is a lack of evidence regarding the potential effects of maternal caffeine, despite these two behaviours being highly correlated. Second, there appears to be an overrepresentation of studies investigating the effect of prenatal smoking and caffeine exposure on externalising problems, compared to internalising problems in offspring. Third, most of the evidence on maternal smoking and caffeine consumption during pregnancy on offspring mental health outcomes is based on observational study designs, which likely suffer from selection, measurement, and confounding biases and therefore cannot differentiate confounded from causal effects. Despite strong theoretical grounds for DNA methylation linking prenatal exposures and offspring mental health problems,

very few studies have formally investigated this research question (Barker, Walton, & Cecil, 2018).

In conclusion, we urgently need better evidence to support public health policies and guidelines around health behaviours during pregnancy. Such policies will be most beneficial to the health of all family members if they are based on robust evidence of causal effects and communicated in a way that avoids blaming mothers. This thesis is aimed at contributing evidence for the establishment of evidence-based guidelines for smoking and caffeine consumption during pregnancy. I am trying to disentangle causal from confounded associations between smoking and caffeine consumption during pregnancy and offspring internalising and externalising mental health outcomes by using genetic variants as proxies for smoking and caffeine consumption during pregnancy. Further, I am investigating offspring DNA methylation as a molecular pathway explaining the role of prenatal exposure to smoking and caffeine on offspring mental health outcomes.

Chapter 2 – Methods and Samples

2.1 Chapter overview

In this chapter, I first introduce the methods that were used in more than one of the analyses of this thesis. The methods that were used just once in this thesis (phenome-wide association study and two-sample Mendelian Randomisation) are discussed in more detail in the individual chapters (Chapters 3 and 6, respectively). The methods described in this chapter include: Mendelian Randomisation (MR), Polygenic Risk Score (PRS) analyses and Epigenome-Wide Association Studies (EWAS).

I then give an overview of the samples that were used for these analyses. Across this thesis, data were analysed from six prospective birth cohort studies, which are all part of the Pregnancy And Childhood Epigenetics (PACE) consortium (Felix et al., 2018): Avon Longitudinal Study of Parents and Children (ALSPAC) (Fraser et al., 2013); Born in Bradford (BiB) (Wright et al., 2013); Generation R Study (Jaddoe et al., 2006), INfancia y Medio Ambiente (INMA) (Guxens et al., 2012); the Norwegian Mother and Child Cohort Study (MoBa) (Magnus et al., 2006), and Etude des Déterminants pré et post natals du développement et de la santé de l'Enfant (EDEN) (Heude et al., 2016).

In addition to describing the methods and samples, I describe publicly available databases, which combine results from studies that have already investigated genetic and epigenetic associations with traits of interest relevant to this thesis, and thus were used to contextualise results from my analyses. These include the genecards database (Stelzer et al., 2016), the GWAS catalog (Buniello et al., 2019), and the EWAS catalog. (http://www.ewascatalog.org/about/).

2.2 Methods

2.2.1 Genetically informed analyses

2.2.1.1 Mendelian Randomisation (MR)

As briefly mentioned in the previous chapter, MR is a useful method to investigate intrauterine effects on long-term outcomes, especially where ethical or practical considerations render randomised controlled trials unfeasible (Davey

Smith, 2008). MR is particularly useful for investigating behaviours that are strongly socially patterned and therefore prone to confounding, such as smoking and caffeine consumption during pregnancy. As genotypes are allocated at conception, reverse causation can generally be ruled out in MR (e.g., it is not possible that starting to smoke as a teenager will influence the assortment of one's own genetic variants). Crucially, the distribution of genetic variants in a population is generally unrelated to confounding variables, due to random segregation and assortment of alleles (Mendel's second law, or law of independent assortment). Therefore, genetic variants can be used as instrumental variables, and the MR design can be thought of as the observational research's natural equivalent to a randomised controlled trial (Davey Smith & Ebrahim, 2003). The simplest possible form of human genetic variation is single-nucleotide-polymorphisms (SNPs), which capture common variation (frequency > 1%) at a single base pair across chromosomes (Davey Smith & Ebrahim, 2003; Hirschhorn & Daly, 2005). Cheap and convenient microarrays, as small as a post stamp, are routinely employed in population-based studies, which can genotype around one million SNPs from samples of DNA collected through saliva, buccal swabs, or blood. These capture most of the common variation in the genome without the need for expensive whole genome sequencing.

The potential of MR to disentangle causal from confounded associations in the context of maternal pregnancy exposures can be illustrated by the example of an MR study of maternal alcohol consumption during pregnancy. Whereas some observational research reported counter-intuitive findings of modest/occasional maternal alcohol consumption during pregnancy being positively associated with cognitive outcomes in offspring, studies using MR found evidence for the opposite effect – i.e., prenatal alcohol exposure being detrimental, which suggests that previously reported observational associations were confounded (Zuccolo et al., 2013). This hypothesis was supported by the observation of moderate alcohol consumption during pregnancy in ALSPAC being associated with higher maternal education, socioeconomic position (SEP) and better nutrition, compared to no or low amounts of alcohol consumption during pregnancy (Zuccolo et al., 2013), as well as a similar effect between paternal alcohol consumption during pregnancy

and offspring intelligence quotient (IQ) in a negative control design (also in ALSPAC) (Alati et al., 2008).

MR produces valid causal estimates of the association between exposure and outcome, under the premise that the instrumental variable assumptions hold, namely that the genetic variants are: (1) strongly associated with the exposure of interest (relevance), (2) not associated with any other variables affecting the outcome (independence, or exchangeability), and (3) only associated with the outcome through the exposure of interest (exclusion restriction) (Davies et al., 2018; Pingault et al., 2018; Taylor, Davies, et al., 2014). Therefore, MR may provide insights into causal effects of smoking and caffeine consumption during pregnancy if: (1) there is evidence that genetic variants are strongly associated with the actual behaviour of smoking or consuming caffeine during pregnancy, (2) there is evidence that the genetic variants for smoking and caffeine consumption during pregnancy are not associated with potential confounding variables, such as SEP, maternal mental health or offspring's own smoking or caffeine consumption, and (3) there is evidence that the genetic variants are only associated with offspring mental health outcomes through their association with maternal smoking or caffeine consumption during pregnancy.

Genetic variants used in MR are commonly selected from GWAS of the exposure of interest – this ensures that the first assumption is met. In a GWAS, SNPs of the entire genome are tested for statistical associations with a trait of interest, which can either be tested by comparing frequencies of alleles between cases and controls, or across a trait's distribution (Marees et al., 2018). Evidence for a genome-wide significant SNP-phenotype association is commonly determined by having a P-value that falls below the genome-wide P-value threshold 5×10^{-08}. Additionally, SNPs can be investigated for indications of violations of assumption two and three, by running additional MR methods that are more robust to pleiotropy (e.g., MR Egger; see 2.2.2 for more information) (Bowden et al., 2015) and checking SNPs for already known associations with other phenotypes through running SNPs through the PhenoScanner (Kamat et al., 2019; Staley et al., 2016).

2.2.1.2 Polygenic Risk Score (PRS) Analyses

For complex traits, such as behaviours during pregnancy and mental health outcomes, the magnitude of the effect sizes of the genome-wide significant SNPs is commonly very small, with SNPs having low predictive power when applied individually (Choi et al., 2020). This led to the assumption that complex traits are most likely polygenic. Instead of a single genetic variant, many genetic variants with small effect sizes contribute to the phenotypic variation that is observed for complex traits (Dudbridge, 2016). A fruitful method to increase predictive power of genetic predictors is to generate a polygenic risk score (PRS) that combines the effects of individual SNPs. The SNPs with the strongest evidence for an association with the trait of interest, commonly determined by the associations with the smallest P-values, are selected for generating the PRS (Choi et al., 2020). The number of SNPs selected for the PRS can vary depending on the P-value threshold chosen and may also include SNPs from GWAS results that have not revealed any genome-wide significant associations. Decreasing the P-value threshold offers the advantage of including a higher number of SNPs in the PRS, which may increase the predictive power of the PRS but at the cost of introducing more noise into the PRS (and thereby decreasing specificity) (Choi et al., 2020). The increase in explained variance, resulting from including SNPs that are only weekly associated with the trait of interest, is unlikely to be large enough to counterbalance the noise that they introduce (Janssens, 2019). Researchers might select their P-value threshold in accordance with their research question. Selecting a more stringent P-value threshold and only including SNPs more likely to be causally associated with the trait of interest, may be useful for research questions that are focussed on causal inference. A more liberal P-value threshold may be useful when the research questions are focussed on prediction. When a stringent P-value threshold is selected, PRS analyses can be used to infer causality by drawing on similar concepts as MR (except of not estimating the causal effect of the modifiable exposure on the disease) (Burgess & Thompson, 2013).

Ideally, the target sample in which one wants to apply the PRS (e.g., for risk prediction or causal inference) is independent of the training/discovery GWAS sample that the PRS is based on (Dudbridge, 2016). A PRS is calculated for each individual in the target sample by summing the number of their risk alleles

weighted by the effect size of the effect allele of the training sample (Dudbridge, 2016). A weighted PRS allows for some SNPs to have a stronger influence on the trait of interest than others (Janssens, 2019). The effect allele is defined as the allele, which has a higher frequency in the population expressing the trait of interest, compared to the population which does not express the trait of interest (e.g., an effect allele for depression would be the allele that is more commonly observed in people with depression than the control group or the effect allele for smoking would be the allele more commonly observed in people who smoke than people who do not smoke). Once the PRS has been generated, a regression analysis can be conducted in the target sample using the PRS as the (continuous) explanatory variable and the trait of interest as the outcome.

In sum, MR and PRS analyses can be used for producing evidence robust to environmental confounding (especially socioeconomic confounding) by using genetic variants as proxies for modifiable exposures. Whereas PRS analyses can be used in an explorative manner to generate new hypotheses (e.g., see Chapter 3) without formal testing of the instrumental variable assumptions, MR is applied to test specific hypotheses, which requires that the instrumental variable assumptions hold in order to yield unbiased results. The next section will discuss threats to the instrumental variable assumptions.

2.2.1.3 Genetic confounding

As mentioned above, using genetic variants as instrumental variables mimics a randomised controlled trial under the premise that genes are allocated at random at conception (Mendel's law of random assortment of genes). This random assortment ensures that participants in the exposed and unexposed group only differ on their level of exposure but are otherwise exchangeable (Pingault et al., 2018). Genetic confounding may threaten the exchangeability assumption. The next section will outline some important potential sources of genetic confounding.

2.2.1.3.1 Population stratification

One form of genetic confounding is population stratification, which occurs when cases and controls systematically differ in their allele frequency because of ancestry differences (Hellwege et al., 2017; Price et al., 2006). Spurious associations may arise when certain populations are more likely to express the

phenotype of interest than others, or because of ascertainment of case/control status. As mating does not occur completely at random but is influenced by geographical proximity, genetic structure is generated in a population (Choi et al., 2020). As risk factors may also be associated with geographical location, population stratification may generate groups with differing levels of risks (Choi et al., 2020). This is problematic for MR, as population stratification implies that exchangeability cannot be ensured, given that the genetic variants were not randomly allocated and therefore people may not have been randomly grouped into "exposed" and "unexposed" groups. A famous example that illustrates the problem of population stratification is that of the "chopstick gene" (Hamer & Sirota, 2000). Due to population stratification, a gene was found to be associated with the use of chopsticks, despite no biological explanation. As the sample consisted of participants of Asian and white European ancestries, who differ in their allele frequencies, a spurious association was created between a certain allele and chopstick use that was due to cultural but not biological reasons (Hamer & Sirota, 2000). If not accounted for, population stratification could threaten the relevance assumption of MR, as the wrong genetic variants might be used as a proxy for the exposure of interest (e.g., if one would have used the genetic variants for chopstick use any effect of the genetic variants on an outcome would have been falsely attributed to chopstick use instead of ancestry). Also, it would have violated the exchangeability assumption, as the exposed and unexposed group would not only differ in terms of exposure (e.g., chopstick use) but also differ in their ancestries (and associated cultural variables). As GWAS and PRS analyses often combine many different populations to meet the sample size needed to detect small genetic effects, population stratification needs to be accounted for. One approach to reduce spurious associations because of population stratification is to conduct analyses separately for different ancestries (e.g., separate analyses for White European and Asian ancestries). An additional routinely used approach to further account for population stratification is to conduct a principal component analysis (PCA). Briefly, in PCA genotype data of the whole sample is used to infer continuous axes that account for as much between-subject genetic variance as possible (Price et al., 2006). These axes, or principal components, can then be added as covariates to the regression models to reduce genetic ancestry differences. The more diverse ancestries a sample

includes, the more principal components are needed to account for population stratification. Generally, the first component explains most of the variance introduced by population stratification, compared to the following ones (Anderson et al., 2010).

2.2.1.3.2 Linkage disequilibrium

Another reason why the exchangeability assumption may not hold within a specific population, is that not all alleles across the genome are allocated completely at random. Linkage disequilibrium (LD) refers to SNPs on a chromosome (particularly those physically close to one another) showing a correlation that is higher than would have been expected by chance, as they are not inherited independently of each other but 'together' on 'LD blocks' (Lawlor et al., 2008; Marees et al., 2018). The implications of LD for MR studies are twofold: On the one hand, LD enables to study a functional genetic variant that has not been assessed, by selecting an available genetic proxy that is in high LD with the functional variant of interest (tagging SNPs). On the other hand, LD may confound MR estimates if the genetic variant used as the instrumental variable is in LD with another variant that independently affects either the outcome (violating the exclusion restriction assumption) or potential confounders (violating the exchangeability assumption) (Lawlor et al., 2008). Pruning and clumping are statistical techniques that can be applied to account for LD. Both methods reduce the number of SNPs within the same genomic region. Pruning removes SNPs at random and clumping removes SNPs based on prespecified LD threshold (Choi et al., 2020).

2.2.2 Pleiotropy

Pleiotropy is defined as a single genetic variant influencing multiple phenotypes and believed to be a common phenomenon in the complex genome-phenome relationship (Hemani, Bowden, et al., 2018; Pingault et al., 2018). There are two forms of pleiotropic associations. One is called horizontal pleiotropy, describing an association between the genetic variant and the outcome that is not explained through the exposure of interest but through independent biological pathways (thus also commonly referred to biological pleiotropy) (Hemani, Bowden, et al., 2018). The second form of pleiotropy is called vertical (or mediated) pleiotropy, where the outcome shows an association with the genetic variant because the

genetic variant is associated with a trait that influences the outcome. Whereas vertical pleiotropy is not problematic for MR because the other trait is mediated through the exposure of interest and therefore does not violate the exclusion restriction assumption (and actually reflects the concept MR is built on), horizontal pleiotropy could lead to biased estimates in MR because it violates the exclusion assumption (Davey Smith & Hemani, 2014). Figure 2.1 illustrates these two forms of pleiotropy with a hypothetical example assessing the effect of the smoking PRS on offspring mental health outcomes. The purple arrows describe the scenario of vertical pleiotropy. Hypothetically, if smoking during pregnancy leads to less social support because of stigmatisation, which in turn negatively influences offspring's mental health development, the intergenerational MR analysis still should yield unbiased results. The pleiotropic path (social support) is associated with offspring mental health outcomes through the trait of interest (maternal smoking) and therewith only indirectly associated with the maternal smoking PRS. The red arrows in Figure 2.1 illustrate a hypothetical scenario for horizontal pleiotropy, which threatens the exclusion restriction assumption. Offspring mental health outcomes are not associated with the trait of interest (maternal smoking) but with the maternal smoking PRS through an independent biological pathway linking the SNPs of the smoking PRS and mental health outcomes (e.g., SNPs of the smoking PRS influencing the hypothalamic-pituitary-adrenal (HPA)-axis and therewith mental health outcomes independently of the actual maternal smoking behaviour). Pleiotropic associations between the PRS for smoking and caffeine consumption and offspring mental health outcomes are investigated in more detail in the next chapter (Chapter 3).

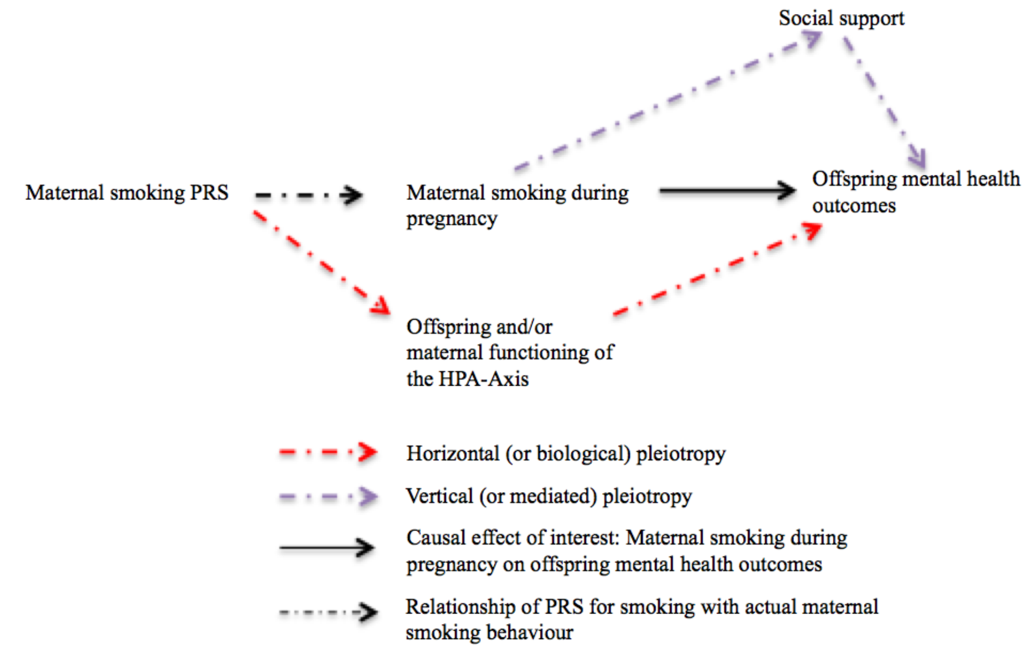


Figure 2.1 Illustration of vertical and horizontal pleiotropy for assessing the effect of maternal smoking on offspring mental health outcomes. *The purple arrows illustrate a vertical pleiotropic path from the maternal smoking PRS to offspring mental health outcomes. The red arrows illustrate horizontal pleiotropic pathway between the maternal smoking PRS and offspring mental health outcomes.*

2.3 Epigenome-wide association studies (EWAS)

As briefly mentioned in Chapter 1, most studies investigating DNA methylation in relation to a phenotype of interest have applied a candidate gene approach. Since the release of Illumina methylation microarrays in 2007 (Bibikova, 2006), which allowed a wider range of DNA methylation sites to be assessed in a more time and cost-efficient manner, EWAS have become the state-of-the-art approach for investigating DNA methylation-phenotype associations (Jones et al., 2018). EWAS allow discovering novel differentially methylated CpG sites associated with a phenotype of interest in a hypothesis-free manner. Further, it is still possible to test previously found DNA methylation-phenotype associations in a candidate-gene hypothesis-testing fashion (if the array covers the candidate regions of interest) (Jones et al., 2018). In the following sections, I briefly discuss how DNA methylation for an EWAS is measured and how an EWAS is conducted.

2.3.1 Assessment of DNA methylation using Illumina microarrays

DNA methylation can be collected from any tissue but in practice is most commonly assessed from saliva, buccal cell swaps, or blood because they can be collected non-invasively (Jones et al., 2018). Different methods can be used to assess DNA methylation, and the most common one is to assess DNA methylation through commercially attainable microarrays from Illumina. Illumina uses sodium bisulfite to detect which CpG sites are methylated (Bibikova, 2006 & Bibikova et al., 2011). Through applying sodium bisulfite to DNA samples, cytosine at unmethylated CpG sites is converted to uracil, whereas methylated CpG sites remain unchanged (Jones et al., 2018). This then allows quantification of the ratio of methylated to unmethylated CpG sites at a given genomic position across cells in form of a β-value. The β-value is commonly calculated using the following formula:

$$\beta = \frac{Max\ (M, 0)}{Max\ (M, 0) + Max(U, 0) + \alpha}$$

In this formula, the nominator represents the maximum DNA methylation signal (M) at a given allele and the denominator represents the sum of the methylated (M) and unmethylated signal (U) plus a constant α, which is recommended to be 100, to stabilize β-values where methylation intensities are small (Bibikova et al., 2011; Du et al., 2010; Weinhold et al., 2016). The resulting β-value is a value between 0 and 1, representing the average percentage (0-100%) of DNA methylation of alleles across all measured cells. To illustrate, a β-value of 0 for a specific CpG would represent that none of the cells in a sample contained DNA that was methylated at that CpG. A β-value of 1 would represent all cells in a sample contained DNA that was methylated at that CpG (both extremes are rarely observed in practice).

Ever since the release of the first Illumina array in 2007 (Bibikova, 2006), the number of CpG sites that can be assessed across the epigenome has tremendously increased, from ~27K to ~800K, representing the fast developments in the field of epigenetics over the past two decades. Still the largest array currently available, which is the EPIC array (Moran et al., 2016), only assesses 853,307 CpG sites, covering less than 5% of CpG sites of the genome. Only one cohort of this thesis

assessed DNA methylation using the EPIC array, of which only common probes with the 450k array were used (see Born in Bradford in section 1.5.4 for more details). Therefore, DNA methylation analyses in this thesis are based on the 450k array (Bibikova et al., 2011), which is discussed in more detail below.

2.3.1.1 The 450K array

The 450k array was released in 2011 and covers approximately 480,000 CpG sites, representing less than 2% of CpG sites of the genome. The selection of probes on the 450k has been based on the co-methylation assumption, which posits that adjacent sites (within 50 bases) are likely to show very similar DNA methylation patterns (Bibikova et al., 2011; Eckhardt et al., 2006) and thus not all sites need to be assessed in order to observe global DNA methylation changes. More detailed information about the selection of probes on the 450k array can be found in the paper by Bibikova and colleagues (Bibikova et al., 2011). Briefly, genomic positions for the 450k array were selected by an epigenetic expert panel, resulting in 482,421 CpG sites, 3,091 non-CpG sites, and 65 randomly selected SNPs. Overall, these probes should be representing DNA methylation at multiple CpG sites (mean = 17.2 probes per gene) of 99% of genes (as annotated by RefSeq) and at 96% of CpG islands and shores (Bibikova et al., 2011). An illustration of DNA methylation patterns at these regions can be found in Figure 2.2 and may help to clarify why DNA methylation at these regions are of particular interest. Most CpG sites in the genome tend to be methylated and are commonly found outside of gene promoter regions. However, CpG islands (Figure 2.2), which are sections of DNA where many successive CpG sites can be found (see blue line labelled CpG density), are typically unmethylated (represented by thin grey line in Figure 2.2) and commonly found in close proximity to promoter sites of 60-70% of all human genes (Illingworth & Bird, 2009). In contrast to CpG islands, DNA methylation adjacent to CpG islands, at CpG shores and shelves, varies greatly (Figure 2.2, grey line). Changes in these DNA methylation patterns are likely to influence gene expression and are therefore an attractive research target. Whereas DNA methylation at CpG sites close to gene promoter sites tends to be negatively associated with gene expression, DNA methylation at gene bodies is more commonly positively associated with gene expression (Jones et al., 2018).

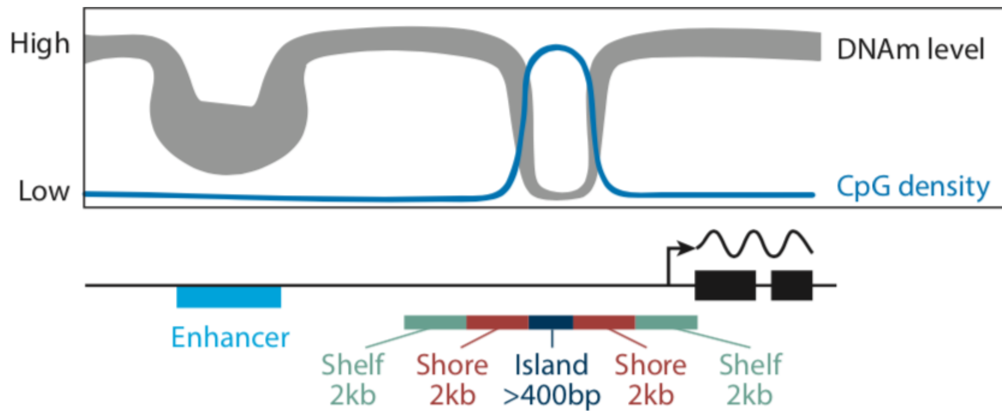


Figure 2.2 Illustration of DNA methylation patterns near a gene region *(credit: Jones et al., 2018). The Black arrow represents the transcription start site and the black rectangles represent a gene. The grey line symbolizes variability in DNA methylation, with a thicker line representing higher variability in DNA methylation levels according to the genomic region. The blue line symbolises CpG density according to genomic region.*

2.3.2 Biological variation in DNA methylation

When conducting an EWAS it is important to be aware of certain biological properties of DNA methylation that need to be accounted for to reduce variation that is not related to the trait of interest. Known biological sources of variation in DNA methylation include tissue-specificity, cell proportion types, age, sex, and population stratification. These sources of biological variation and how to adjust for them are discussed in more detail below.

2.3.2.1 Tissue specificity

DNA methylation is tissue specific, meaning that DNA methylation levels, observed at the same genomic position in a cell, are likely to differ across different tissue types. It is therefore unclear whether DNA methylation assessed in one type of tissue is able to function as a surrogate for DNA methylation in another tissue (Lokk et al., 2014). Whereas tissue-specificity might be less problematic for analyses trying to establish DNA methylation in an accessible tissue (such as blood or saliva) as a biomarker for a phenotype of interest, tissue-specificity is highly relevant when the DNA methylation analysis is focussed on understanding biological pathways to the phenotype of interest. Regardless of the focus of the EWAS, studies investigating the association between DNA methylation and psychological/psychiatric phenotypes commonly rely on the assumption of cross-tissue concordance because of the inability to assess DNA

methylation in the tissue of most interest, which is human brain tissue (Walton et al., 2019). Efforts have been made to test the cross-tissue concordance assumption between peripheral and brain tissue derived DNA-methylation. Initial findings looked promising, with studies that applied a between-subjects approach reporting high correlations between DNA methylation in peripheral tissue and brain tissue. Across-subject correlations of DNA methylation in post-mortem brain tissue were found to be ranging from 0.77-0.91 with blood (Braun et al., 2019; Davies et al., 2012; Horvath et al., 2012), 0.90 with saliva, and 0.85 with buccal cell tissue (Braun et al., 2019). However, when considering both, between- as well as within-subject correlations of DNA methylation levels across different tissues, correlations tend to be much smaller. Evidence for within subjects correlations of blood-brain DNA methylation was only found for 1.4% to 7.9% of CpG sites (Hannon et al., 2015; Walton et al., 2016). Further, a study by Edgar and colleagues (Edgar et al., 2017) calculated correlations between DNA methylation in post-mortem brain tissue and paired blood DNA methylation levels of 16 subjects. They found that blood-brain DNA methylation at most CpG sites within subjects did not correlate and those CpG sites which did show evidence for a correlation, showed insufficient variability between individuals, indicating low susceptibility of those sites to environmental influences (and therewith low relevance for EWAS). Those CpG sites which met sufficient inter-individual variability to indicate susceptibility to environmental influences, showed a within-subject blood-brain concordance ranging from 0.33-0.40, which still indicates substantial variability between blood-brain DNA methylation. Likewise, cross-tissue comparisons of DNA methylation using brain tissue from living subjects obtained during neurosurgery for epilepsy, found that 26.6% of CpG sites in blood, 17.6% in saliva, and 20.4% in buccal cells showed sufficient inter- as well as intra-individual variability and were correlated with DNA methylation levels in brain tissue (Braun et al., 2019). Similar to previous studies, Braun and colleagues (2019) also found higher between-subjects correlations for brain DNA methylation and saliva ($r = 0.90$), blood ($r = 0.86$) and buccal cell DNA methylation ($r = 0.86$) than the corresponding within-subjects correlations (brain-saliva: mean $r = 0.12$; blood-brain: mean $r = 0.15$; brain-buccal cell: mean $r = 0.14$).

Lack of overlap between DNA methylation in blood and post-mortem brain tissue may be due to changes in brain DNA methylation related to death, or the cause of death, as well as to commonly limited sample sizes because of the scarcity of brain tissue for research purposes. In addition to studies using post-mortem brain tissue, some studies were able to test concordance with brain tissue from living subjects during brain surgery (Braun et al., 2019; Walton et al., 2016). Findings from these studies are in line with findings using post-mortem brain tissue. It is noteworthy that these studies show limited power due to restricted sample sizes and were collected from patients receiving neurosurgery because of severe epilepsy, which may also have impacted DNA methylation.

In sum, tissue-specificity needs to be considered when inferring biological pathways from EWAS using peripheral tissue as surrogates for brain tissue. When interested in intra- and inter-individual correlations of brain tissue with peripheral tissue, peripheral blood seems to be the best available surrogate for brain tissue (slightly higher concordance than saliva or buccal when considering within-subject variation) (Braun et al., 2019). No data are available on the overall concordance between DNA methylation in cord blood and brain tissue, yet specific CpG sites showing evidence for DNA methylation changes in cord blood can be looked up on databases of peripheral blood-brain tissue concordance (Edgar et al., 2017), to get an indication for whether these DNA methylation changes are likely to be mirrored in brain-tissue (Caramaschi et al., 2017, 2020).

2.3.2.2 Estimated cell proportions

In addition to cross-tissue variation, there is also substantial variation in DNA methylation within the same tissue caused by the different cell-types it contains (Houseman et al., 2012; Jones et al., 2018). Different cell types might confound or mediate associations between DNA methylation and a phenotype of interest, especially if the phenotype of interest also shows cell-type associated variation (Yousefi, Huen, Quach, et al., 2015). Therefore, it is often recommended to adjust for cell type proportions in EWAS and/or to explore associations between cell type proportions and the phenotype of interest. Cell type proportions can either be measured directly in each sample or be estimated. Due to time, cost and laboratory storage restrictions, and because many studies rely on pre-extracted DNA samples and do not have access to original samples, the latter is the more

common approach. A popular method to estimate cell proportions in blood is the Houseman method (Houseman et al., 2012), which builds a model, based on data of cell-types of a reference panel that can estimate cell proportions in the target sample. In fact, it is the differences in DNA methylation patterns between the cells, as established in the reference data that help to distinguish the proportions of each cell type in the target sample. The estimated proportion of each cell type can then be added as covariates to the EWAS models. As recent studies have shown that cell proportions are influenced by age (Jaffee & Price, 2008; Yousefi, Huen, Quach, et al., 2015) and most reference panels are based on adult populations, cord blood reference panels have been established that enable accurate prediction of cell proportions in cord blood (Bakulski et al., 2016; Gervin et al., 2016, 2019). For some other tissues than blood, such as saliva and placental tissue, the underlying cell type proportions may not have been clearly established in reference data and thus a reference free method may be more appropriate (Houseman et al., 2014; McGregor et al., 2016). However, reference-free analyses may come with the cost of reducing potentially interesting phenotypic variation (Gervin et al., 2016; McGregor et al., 2016). All EWAS analyses in this thesis used the Houseman method to estimate cell proportions in blood (reference panel for cell proportions at birth: Gervin et al., 2016; reference panel for cell proportions in childhood: Reinius et al., 2012). In addition to the hypothesis of environmental exposures leading to changes in DNA methylation between exposed and non-exposed individuals, the hypothesis exists that exposures may influence cell type proportions (Lappalainen & Greally, 2017). Therefore, I also tested the association between the phenotypes and estimated cell proportions in my analyses.

2.3.2.3 Age

Many studies have shown associations between chronological age and DNA methylation at various developmental stages (Li et al., 2018; Mulder et al., 2020; Teschendorff, West, et al., 2013; van Dongen et al., 2016; Xu et al., 2017). Overall, it is suggested that DNA methylation is more variable early compared to later in development (Mulder et al., 2020; Reynolds et al., 2020). Age dependent epigenetic changes, also referred to as epigenetic drift (Teschendorff, West, et al., 2013), were initially observed in twin studies, which found higher variability of

DNA methylation with increasing age. Whereas monozygotic twins showed an almost identical DNA methylation profile in early life, epigenetic differences increased with increasing age (Fraga, Ballestar, Paz, Ropero, Setien, Ballestar, Heine-Suñer, Cigudosa, Urioste, Benitez, et al., 2005; Talens et al., 2012). A recent meta-analysis of two large birth cohorts (N = 2,348) has investigated DNA methylation changes from birth to adolescence and found that about half of the assessed CpG sites linearly change over development, specifically in the first decade of life, and generally show a decrease in DNA methylation (Mulder et al., 2020). In addition to chronological age, gestational age was found to be strongly associated with cord blood DNA methylation. A meta-analysis of 17 birth cohorts has found DNA methylation at over 8,000 CpG sites to be associated with gestational age, yet of these only one CpG site showed consistent evidence for an association with DNA methylation in childhood and adolescence (Merid et al., 2020). Yet, 280 of the gestational-associated CpG sites were in common with a list of CpG sites found to be associated with childhood age (Xu et al., 2017). Alternative explanations for lack of consistency across development in Merid and colleagues' (Merid et al., 2020) meta-analysis may be the smaller sample sizes at later DNA methylation assessments or the increased number of other postnatal environmental exposures that have distorted the effect of gestational age (Merid et al., 2020). Noteworthy, the cord blood CpG sites associated with gestational age showed a higher variability in DNA methylation in early childhood than the non-gestational age associated CpG sites, indicating that these CpG sites may be specifically prone to changes in DNA methylation. Due to the influence of age on DNA methylation it is recommended to adjust for age in EWAS analyses (Jones et al., 2018).

2.3.2.4 Sex

There are substantial differences in DNA methylation patterns between female and male sex. DNA methylation is involved in the deactivation of the second X chromosome in females. Due to the profound differences in DNA methylation at the sex chromosomes, these are typically analysed separately or removed from the analyses (Jones et al., 2018). Even when restricting analyses to autosomes and removing probes on autosomes, which are cross-reactive with sex chromosomes, sex-specific DNA methylation differences can still be observed (Yousefi, Huen,

Davé, et al., 2015). The mechanism for these sex-specific differences in DNA methylation is not clear, however, it is advised to either adjust for sex or to inspect sex-specific associations between DNA methylation and the phenotype of interest through stratifying by or including sex as a moderator (Jones et al., 2018).

2.3.2.5 Population stratification

As previously mentioned, population stratification may be a confounding factor in genetic analyses and is therefore commonly adjusted for by stratifying analyses by ancestry (Hellwege et al., 2017). Whereas PCA is commonly applied to adjust for genetic variation in the data attributable to genetic ancestry, stratification of ancestry is commonly based on self-reported ethnicity. Self-reported ethnicity is a complex multifaceted construct that may reflect genetic and/or cultural aspects (Barnholtz-Sloan et al., 2008). Ancestry differences, specifically if assessed through self-report, may also be associated with phenotypic differences, such as differences in behavioural, societal, and cultural aspects (Barnholtz-Sloan et al., 2008). As environmental and genetic effects influence DNA methylation, ancestry may be an important cofounding factor in EWAS analyses (Fagny et al., 2015; Galanter et al., 2017). For instance, a study investigating DNA methylation and genotype data of 573 children with varying Hispanic origin found evidence for genetic, but even stronger evidence for self-reported ancestry, being associated with DNA methylation differences (Galanter et al., 2017). Whereas genetic ancestry, determined by PCA, was associated with 194 differentially methylated CpG sites, self-reported ancestry was associated with 916 CpG sites. Even after adjusting associations between self-reported ethnicity and DNA methylation for genetic ancestry, still 34% of variance remained, indicating an influence of self-reported ethnicity on DNA methylation that is not explained by population stratification. The self-reported ethnicity associated CpG sites were enriched for CpG sites that have previously been associated with phenotypes of socioeconomic position, including exposure to maternal smoking during pregnancy, car exhaust gas and psychosocial stress. First, this illustrates the importance of accounting for self-reported ethnicity differences in samples combining populations with diverse ethnic identities. Second, it stresses the need for researchers to explicitly define how ancestry/ethnicity has been assessed in their study to ensure that adequate inferences can be drawn (Ali-Khan et al., 2011; Barnholtz-Sloan et al., 2008). In

this thesis, the only DNA methylation sample with large proportions of diverse ethnicities that was used was the Born in Bradford cohort (BiB; see 2.6.4 for cohort description) and thus the EWAS analysis of this cohort was stratified by self-reported ethnicity.

2.3.3 Technical variation in DNA methylation assessment

In addition to biological sources of variation in DNA methylation, another source of unwanted variation might be introduced by technical handling of DNA methylation samples. In this section, the sources of technical variation and how to account for them are discussed.

As samples of DNA methylation are collected in different batches, a large source of technical variation may stem from batch effects. Some sources for batch effects can be controlled for at the study design stage by closely protocolling difference in sample collection, such as dates on which the individual batches were collected and the technical personnel that conducted the assessment. Further, randomising samples to different batches (e.g., avoiding assessing all cases and controls in the same batch) might help to reduce batch effects. Additional types of technical variation can be controlled through 848 control probes that were intentionally included on the 450k array to assess non-biological variation of the data (Fortin et al., 2014). Functional normalisation can be applied to the DNA methylation data using the variation assessed by the control probes to disentangle technical from biological variation (Fortin et al., 2014; Min et al., 2018). In addition, a selection of probes of the 450k microarray, summarised by Chen et al. (Chen et al., 2013), have consistently been classified as problematic and thus should be removed from the DNA methylation samples. Also, poor quality probes can be determined by the calculation of a detection P-value, which determines, based on the signal captured by the control probes on the array, whether the probe was able to capture a true signal or noise (Heiss & Just, 2019). The P-value cut-off for determining whether a probe captured a high-quality signal is chosen by the researcher. Generally, probes with detection P-value larger than 0.05 or 0.01 are removed prior to the EWAS analysis (Heiss & Just, 2019). In the EWAS analyses of this thesis, probes were removed if 5% of the samples had a detection P-value > 0.05. Another quality assessment for the DNA methylation data is to remove probes that are statistically classified as outliers. The Tukey method (Tukey, 1977) was

used to identify outliers in the EWAS analyses of this thesis, where probes are removed prior to the EWAS analysis that fall outside 3 times the interquartile range (Sharp, Arathimos, et al., 2018; Sharp et al., 2021).

In order to account for any remaining unmeasured sources of non-biological variation, surrogate variable analysis (SVA) was applied (Leek et al., 2012). SVA helps to remove variance of unknown latent variables underlying the data and thus can be applied to reduce batch effects in microarray data (Leek et al., 2012).

2.3.4 Statistical analysis

The statistical analysis for the meta-EWAS conducted as part of this thesis followed the pipeline for meta-EWAS analyses used by Sharp and colleagues (Sharp et al., 2021). The pipeline includes nine quality control and analysis steps that could be conducted in EWAS meta-analyses. The first six steps were applicable to the analyses of my thesis (Figure 2.3). Step one includes proper filtering of probes to remove technical and biological variation as outlined above. The next steps will be discussed in detail in the following sections.

1. Filter probes from cohort results files to remove probes that are not common to both the EPIC and 450k array, control and QC probes, probes on SNPs, cross-hybridizing probes according to Chen *et al.*[57] and probes on the sex chromosomes.
2. Perform quality checks of cohort results by plotting correlation matrices of effect estimates generated by different models, generating QQ plots and calculating Lambda values, plotting the distribution of effect estimates and producing 'precision plots' of 1/median standard error against the square root of the sample size for each cohort and model.
3. Conduct a fixed-effects meta-analysis using METAL for each model.
4. Adjust meta-analysis *P*-values for multiple testing using the false-discovery-rate (FDR) method. The threshold used to define statistical evidence of an association was an FDR-adjusted *P*-value <0.05.
5. Perform checks of meta-analysed results by plotting a correlation matrix of effect estimates generated by different models, generating QQ plots and calculating Lambda values.
6. Conduct a leave-one-out analysis using the R package metafor[55] at sites with the smallest *P*-values. The leave-one-out survival criteria we specified are: when any single cohort is omitted, the meta-analysis effect estimate should be in the same direction, not attenuate substantially (arbitrarily defined as >20% change-in-estimate) and not have a confidence interval that crosses the null.

Figure 2.3 First six steps of the EWAS meta-analysis pipeline introduced by Sharp and colleagues. *Only steps 1-6 were applicable to the analyses of this thesis. Steps 7-9 were specific to the research design of the study by Sharp and colleagues (credit: Sharp et al., 2021).*

2.3.4.1 Probe-level analysis

As DNA methylation is assessed as a continuous score, which in theory ranges from 0-100%, the degree of association between DNA methylation and phenotypic variation can be estimated through linear regression analyses. Depending on the specific research question, changes in DNA methylation can be investigated as the exposure or outcome. Processed DNA methylation is commonly stored in the form of a matrix, with rows representing the number of assessed probes (approx. 480,000) and columns representing participants within the sample (N columns = N participants). The regression analysis is performed in a row-wise manner, so that, for each CpG, participants' DNA methylation values are compared to their values of the trait of interest (Ritchie et al., 2015). In this thesis, I use DNA methylation as the outcome in my studies of associations with maternal caffeine (Chapter 4) as well as the exposure in my studies of associations with childhood internalising problems (Chapter 5). When using DNA methylation as the exposure, the resulting regression coefficient represents the change in the trait of interest that is associated with a one-unit change in DNA methylation. As a one-unit change in DNA methylation represents a complete change from completely unmethylated (methylation beta value = 0; 0%) to completely methylated (methylation beta value = 1; 100%), across all measured cells, it is more sensible to divide the regression coefficient by 10, to receive a more biologically plausible coefficient. The coefficient would then represent the change in the trait of interest that is associated with an increase in percentage DNA methylation of 10% at a given CpG site. When DNA methylation is modelled as the outcome, the regression coefficient represents the change in DNA methylation at a given CpG site that is associated with a one-unit increase in the trait of interest. Regardless of whether DNA methylation was used as the exposure or outcome, each EWAS analysis in this thesis included a crude model, only adjusted for estimated cell proportions, a covariate model that included all other covariates, and two covariate adjusted models stratified by offspring's sex. All individual cohort analyses were performed in R (R Core Team, 2014). Summary results of the probe-level results were prepared for meta-analysis using R.

2.3.4.2 Quality control checks for the probe-level analyses (Step 2 of EWAS meta-analysis pipeline)

2.3.4.2.1 Cohort quality control checks

Once the EWAS analysis has been conducted there are several quality control checks that can be performed to reassure that there were no problems with the data or the conduction of the analysis. The quality control checks applied in this thesis follow the EWAS meta-analysis pipeline, which has been proposed by Sharp and colleagues (Sharp et al., 2021). First, the correlation between the regression coefficients of each model were calculated and visualised in form of a correlation plot. The correlation plots serve as a form of sanity check to ensure that estimates of similar models are roughly correlated, and models are estimating the same effect. Second, the P-value distribution of the regression coefficients of each model was inspected through the creation of Quantile-Quantile (QQ)-plots. In a QQ-plot, the observed P-value distribution is compared to the distribution that would be expected under the null hypothesis (Ehret, 2010). If some P-values are smaller than would be expected under the null hypothesis, some dots at the end will come off the null line (P-values are sorted from largest to smallest). If dots systematically come off the line at lower values on the x-axis, caution is warranted, as it might represent systematic confounding, for instance through population stratification (Ehret, 2010). Visual inspection of the P-value distribution through QQ-plots can be supported through the calculation of the genomic inflation factor, Lambda (λ). A λ-value of 1 indicates that the observed P-value distribution is identical with the expected null-distribution, indicating no effect or confounding. Whereas we would expect a slightly larger λ-value than 1 (e.g., 1.04) if there were some significant associations between DNA methylation at certain CpG sites and our trait of interest, a largely inflated lambda value is indicative of systematic differences likely due to confounding (Yang et al., 2011). Third, for each model precision plots were generated that visualise the relation of the average statistical precision of the EWAS results and the sample size of each cohort (Figure 2.4). The x-axis represents the square root of the sample size and the y-axis represents 1/median standard error (Figure 2.4). The diagonal line represents the expected precision given the sample size, and strong deviations

from the line could be indicative of issues with the analysis and/or data for that cohort (Figure 2.4).

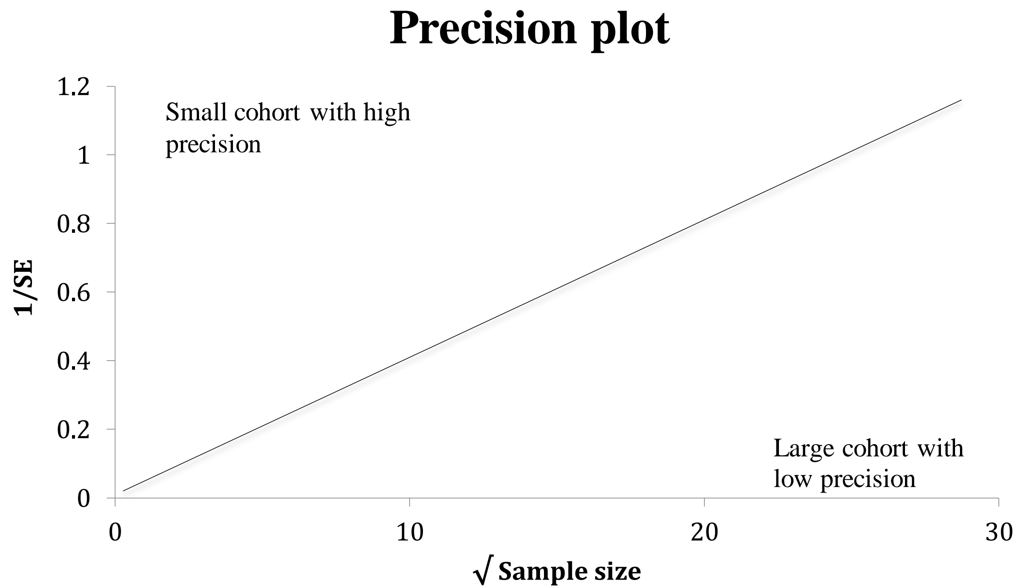


Figure 2.4 Precision plots for cohort results. *The diagonal line represents expected precision according to sample size.*

2.3.4.3 Meta-analysis of probe-level cohort results (Step 3 of EWAS meta-analysis pipeline)

The prepared summary results were meta-analysed with fixed effect estimates weighted by the inverse of the variance using the software METAL (Willer et al., 2010). A fixed effects model was chosen as it assumes that the cohorts stem from the same overarching population and thus the effects should be the same across studies (Borenstein et al., 2007). The only error variance that is considered in a fixed-effects model is the within-cohort measurement error. The observed effect *T* in a cohort *i*, referred to as T_i, results from the sum of the common effect μ and the within-cohort error, ε_i:

$$T_i = \mu + \varepsilon_i$$

The weights ensure that a higher weight is attributed to cohorts with more precise estimates (Borenstein et al., 2007). The weights are calculated according to:

$$w_i = \frac{1}{v_i}$$

Where w_i is the weight for cohort i, which is calculated by dividing 1 by the variance (v) of cohort i. The estimated meta-analysed effect, $\bar{T}$, is calculated by first, summing up the products of each cohort's effect size and their corresponding weight (nominator) and second, dividing it by the sum of the weights (denominator):

$$\bar{T} = \frac{\sum_{i=1}^{k} w_i T_i}{\sum_{i=1}^{k} w_i}$$

To obtain the standard errors (SE) of the estimated weighted effect, $\bar{T}$, the square root of the inverse of the sum of the weights is taken (Borenstein et al., 2007):

$$SE(\bar{T}) = \sqrt{\frac{1}{\sum_{i=1}^{k} w_i}}$$

2.3.5 Meta-analysis quality control checks (Step 5 and 6 of meta-EWAS pipeline)

Quality checks of the meta-analysed results were conducted in a similar vein as for the individual cohort's probe-level results. Correlation matrices of effect estimates of each model and QQ-plots and Lambda values of the P-values of the meta-analysed coefficients were generated. In addition, a leave-one-out analysis was conducted to investigate whether one of the cohorts may be driving the meta-analysed effects. As the name already indicates, a leave-one-cohort-out analysis inspects how the meta-analysed results change if each cohort is sequentially omitted from the meta-analysis. If meta-analysed results significantly change with the removal of one cohort, this provides indication for that cohort exerting a dominant effect on the meta-analysed results that is observed when including all cohorts. As proposed by Sharp and colleagues (Sharp et al., 2021), for this thesis a critical change in results was determined by a change of direction in the meta effect estimate, a more than 20% attenuation to the null, or a confidence interval that includes zero after omission of a single cohort.

2.3.6 Differentially methylated regions

A complementary analysis to the probe-level analysis is a differentially methylated regions (DMR) analysis. In contrast to the probe level analysis, which investigates associations between single probes and a trait of interest, a DMR analysis takes into account groups of CpG sites that are adjacent and show a similar direction of association with the trait of interest (Jones et al., 2018). One of the main advantages of DMR analyses is that they require fewer tests than probe-level analyses and therefore enhance the statistical power in commonly underpowered EWAS (Suderman et al., 2018). Furthermore, DMR analyses are considered to be more biologically meaningful as there is evidence that adjacent CpG sites underlie the same genomic control and are involved in similar biological functions (Eckhardt et al., 2006; Kuan & Chiang, 2012; Saffari et al., 2018). One drawback of DMR analyses is that the magnitude of resulting effect estimates cannot be interpreted biologically, as is possible for a probe-level analysis. The DMR effect size is a meta-analysed coefficient across multiple CpG sites contained within the DMR, each with a unique effect and variance. However, if a DMR analysis is conducted complementary to a probe-level analysis, the probe-level analysis can be consulted to get an indication for the effect size. The DMR analyses in this thesis were conducted using the R-package dmrff (Suderman et al., 2018).

2.4 Mendelian Randomization applied to epigenetic data

As outlined above, DNA methylation is dynamic and thus as likely to suffer from reverse causation (less of a problem when using DNA methylation assessed in cord-blood because of the temporal order of events: offspring mental health outcomes cannot have influenced their DNA methylation levels at birth) and confounding (a factor that affects depression and methylation may explain any depression DNA methylation-association). These problems can be reduced by applying MR to inspect causal effects of DNA methylation on a phenotype of interest, or vice versa. When DNA methylation is used as the outcome, SNPs from GWAS of the exposure of interest can be used as a genetic proxy. When DNA methylation is studied as the exposure, SNPs that proxy for DNA methylation at CpG sites associated with an exposure of interest (e.g., caffeine or smoking), identified from EWAS, can be used to inspect the causal relationship

(Relton & Davey Smith, 2012; Richardson et al., 2018). As for other exposure phenotypes, SNPs discovered from GWAS are commonly selected as instruments that proxy for DNA methylation. SNPs that associate with DNA methylation variation at a CpG site are called quantitative trait loci (mQTL). If the mQTL is in proximity to the CpG site (usually within 1 Mb of each side of the CpG site) it is called a cis-SNP or cis-mQTL (Min et al., 2020). If the mQTL is distal (> 1Mb away from each side of the CpG site) to the CpG site or even on another chromosome, it is commonly referred to as trans-SNP or trans-mQTL (Min et al., 2020). Cis-SNPs are more commonly used as instruments for DNA methylation than trans-SNPs, as most mQTLs are cis (~91%) and have a larger effect on DNA methylation levels than trans-SNPs (Gaunt et al., 2016; Min et al., 2020). Adding to this, trans-SNPs tend to be more prone to horizontal pleiotropy (Hemani, Bowden & Davey Smith, 2018). As in conventional MR, the instrumental variable assumptions must hold to derive accurate conclusions about the effect of DNA methylation on an outcome. When MR is applied to a DNA methylation context, the assumptions are as follows: First, the cis-SNP must be robustly associated with DNA methylation at the CpG site of interest. Second, the cis-SNP should only be associated with the phenotype of interest through exposure-associated DNA methylation variation. Third, the cis-SNP should not be associated with any confounding variables of the cis-SNP-outcome relationship. The first MR assumption should be met by selecting genome-wide significant cis-SNPs, which were replicated in independent samples (Hemani et al., 2018). As mentioned before, the second and third assumption cannot be tested directly and are threatened by horizontal pleiotropy (the cis-SNP is not only associated with DNA methylation at a given CpG site but also independently associated with other phenotypes that may confound associations). Risk of pleiotropy may be reduced by selecting cis- instead of trans-SNPs and by running sensitivity analyses using methods that are more robust to pleiotropy, such as MR-Egger, weighted median, and weighted mode MR (Hemani et al., 2018). However, pleiotropy detection using these methods is limited by the fact that commonly only few independent mQTLs can be identified per CpG site (Jamieson et al., 2020). If MR results indicate evidence in support of DNA methylation levels influencing the phenotype of interest, colocalization methods may be applied to provide evidence

for causation by investigating if the same genetic variant is affecting variation in DNA methylation and the phenotype of interest (Richardson et al., 2018).

2.4.1.1 cis-SNP selection

Large databases have been made available to identify adequate genetic instruments that can be used as proxies for DNA methylation at CpG sites. The mQTL database http://www.mqtldb.org (Gaunt et al., 2016), is a catalogue of SNP-DNA methylation associations (based on the 450k array) within the Accessible Resource for Integrated Epigenomic Studies (ARIES) sample (N = 1,018; Relton, Gaunt, et al., 2015; see sample description for more details about ARIES). Recently, a new data catalogue of SNP-DNA methylation associations called *the genetics of DNA methylation consortium* (GoDMC) has been published. The GoDMC is a large collaboration project between more than 50 research groups that contribute data to investigate the genetic contribution to DNA methylation variation (http://www.godmc.org.uk/) (Min et al., 2020). The GoDMC has investigated the association between genotype and DNA methylation variation using ~10 million genotypes from the 1,000 genome project and blood DNA methylation data from ~30,000 European participants (Min et al., 2020) and is therefore much higher powered than the ARIES based mQTL database (N = 1,018, http://www.mqtldb.org) (Gaunt et al., 2016). The results are now available on a database, containing around 300,000 independent mQTL loci (http://mqtldb.godmc.org.uk/). In this thesis, I use the GoDMC database in Chapter 6 to select genetic instruments as proxies for local DNA methylation associated with prenatal smoking and caffeine exposure.

2.4.2 One-sample and two-sample MR

Depending on the availability of data, different MR approaches may be applied to investigate a causal effect of DNA methylation on a phenotype of interest. If a single sample has genetic, DNA methylation, as well as data on the phenotype of interest, a one-sample MR approach may be applied (Richardson et al., 2018). If a single sample cannot meet these data requirements, a two-sample MR approach may be applied instead. The latter approach approach has the advantage to obtain the data needed from two different samples, one with genetic data and DNA methylation (not requiring data of the phenotype of interest; Figure 2.5 red rectangle) and one with genetic data and the phenotype of interest (not requiring

DNA methylation data; Figure 2.5 blue rectangle; Richardson et al., 2018). The two-sample MR approach tends to be higher statistically powered, as it allows using summary data from large GWAS studies. *MR-base* (https://www.mrbase.org/) is a platform that enables to apply a variety of MR methodologies to curated and harmonised GWAS summary statistics in a user-friendly manner (Hemani, Zheng, et al., 2018). The database includes summary statistics of over 1,673 GWAS and provides all reported SNP-phenotype associations, without restricting to statistically significant results (Hemani, Zheng, et al., 2018). Two-sample MR using MR base was applied in Chapter 6 of this thesis to investigate a potential causal effect of prenatal smoking and caffeine exposure associated DNA methylation on offspring mental health problems.

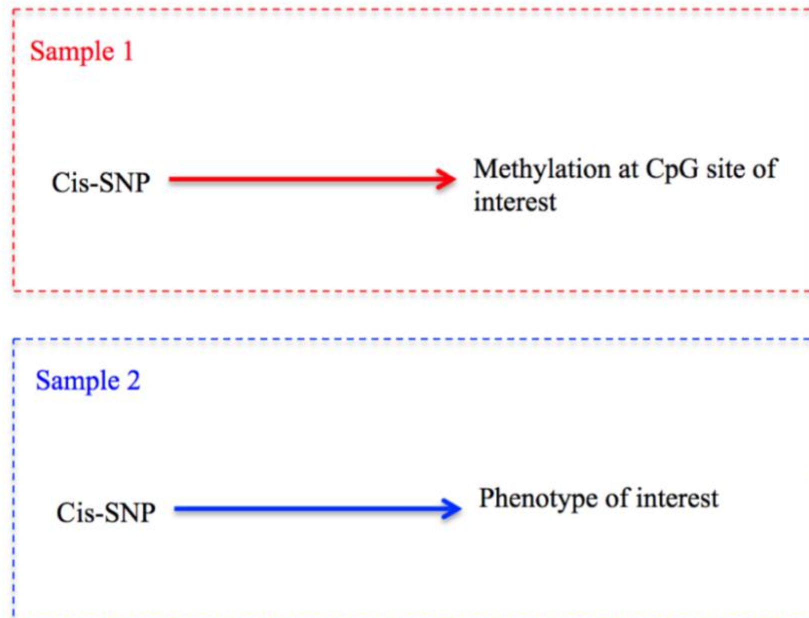


Figure 2.5. Illustration of a two-sample MR analysis investigating the causal effect of exposure-associated DNA methylation on a phenotype. *Using data of sample 1 (top rectangle in red) the association between a genetic variant on DNA methylation at a selected CpG site is estimated. Using data of sample 2 (bottom rectangle in blue) the association between a genetic variant that is a proxy for DNA methylation changes at a selected CpG site and the phenotype of interest is estimated.*

2.5 *Multiple-testing correction*

Omics studies, including EWAS, involve conducting many statistical tests. By chance alone, we would expect a proportion of those statistical tests to produce low P-values. Analyses that run multiple tests on a signal dataset have an inflated chance to tap into Type I error (the probability that the null hypothesis is rejected even though it is true), leading to an increased chance of false positive associations (Shaffer, 1995). Whereas a commonly chosen risk threshold for a false positive finding in statistical analyses is 5% (there is a 5% chance), this risk increases exponentially with the number of tests that are run using the same data set (Goldman, 2008; Shaffer, 1995). As shown in the formula below, if running 20 independent tests on the same data set, using a probability threshold of 5%, the probability (P) for false positive finding increases from a 5% chance for one test, to a 64% chance when running 20 tests (Goldman, 2008):

$$
\begin{aligned}
P\ (at\ least\ one\ false\ positive\ finding) &= 1 - \mathrm{P}\ (no\ false\ positive\ finding) \\
&= 1 - (1 - 0.05)^{20} \\
&= 0.64
\end{aligned}
$$

Two common methods to account for the number of tests that (epi)genetic analyses require are the Bonferroni correction (Shaffer, 1995) and the Benjamin-Hochberg False Discovery Rate (FDR) (Benjamini & Hochberg, 1995). The Bonferroni correction adjusts for the inflated Type I error by dividing the P-value threshold (α) that would be applied to determine statistical significance of a single test (often 0.05) by the overall number of independent tests conducted with the same data (n) (Shaffer, 1995):

$$\alpha_{BF} = \frac{\alpha}{n}$$

The Bonferroni correction is considered to be overly stringent for Type I adjustments in (epi)genetic studies as there is good evidence that the conducted tests are not independent of each other, and thus correction sacrifices unreasonable amounts of statistical power (Shaffer, 1995). As mentioned earlier, SNPs tested in genetic association studies are likely to be subject to linkage disequilibrium, where SNPs on the same chromosome show associations that are

higher than would have been expected by chance (Marees et al., 2018). Likewise, CpG sites, which are in close proximity to another, are likely to be correlated (see section 1.3.6 differentially methylated regions) (Saffari et al., 2018). Multiple-testing correction using FDR also assumes independence between tests but is less stringent than the BF method (Benjamini et al., 2001). When applying FDR, each test's P-value is adjusted based on a set number of false-positives (Benjamini & Hochberg, 1995). As with α, it is up to the researcher to decide on the expected proportion of false discoveries (Benjamini et al., 2001) and a popular approach is to expect a 5% FDR (Marees et al., 2018; Sharp, Arathimos, et al., 2018; Sharp et al., 2021). The FDR is defined as:

$$FDR = \frac{Expected\ Nfalse\ positives}{Expected\ N\ false\ positives + Expected\ N\ true\ positives}$$

2.6 Cohorts

The next section will briefly discuss the cohort data that was used in the analyses of this thesis. ALSPAC was used in all analyses and is therefore discussed in most detail, followed by Generation R and MoBa, which were used for the analyses in Chapter 4 and 5. BiB, INMA, and EDEN only contributed data to the analyses of Chapter 4 and will thus only be discussed briefly.

2.6.1 Avon Longitudinal Study of Parents and Children (ALSPAC)

2.6.1.1 Overview

ALSPAC is a multigenerational cohort study that started recruiting pregnant women in the beginning of the 90's in the greater Bristol area, in South-West England (Boyd et al., 2013; Fraser et al., 2013). ALSPAC is therefore also known as the "Children of the 90's study". Data collection is still on-going and recently also started collecting data on a second generation called "Children of the Children of the 90's" (CoCo90s) (Lawlor, Lewcock, et al., 2019). The data used in this thesis only focuses on the first generation of mothers (G0) and their children (G1), which included 14,541 pregnancies. The motivation to initiate ALSPAC was to increase understanding of modifiable exposures on development and health. In order to pursue this goal, a wealth of data was collected including genetic, epigenetic, biological, psychological, social, and other exposure phenotypes (Boyd et al., 2013). The ALSPAC study was approved by the

ALSPAC Ethics and Law Committee and the Local Research Ethics Committees and informed consent for the use of data collected via questionnaires and clinics was obtained from participants. An overview over the variables that have been collected can be found in the data dictionary and variable catalogue, which can be downloaded from the ALSPAC website (http://www.bristol.ac.uk/alspac/researchers/our-data/). The variable catalogue enables to search for specific variables of interest and can be found at http://variables.alspac.bris.ac.uk/.

All the analyses in this thesis used data from ALSPAC. In Chapter 3, genotype and phenotype data of mothers of G0 and children of G1 were used. Chapter 4 used maternal and offspring genotype data, as well as maternal and paternal phenotype data and cord blood DNA methylation from their children (G1). In Chapter 5, phenotype data from mothers of G0 and their children (G1) were used, as well as DNA methylation at birth and childhood of G1 children. As this thesis only includes G0 parents and their children (G1) the description of participants and data only focusses on these generations of ALSPAC.

2.6.1.2 Participants and data collection

As many pregnant women as possible, living in the Bristol area (formerly the county of Avon) with an expected delivery date that fell in between April 1991 and the end of 1992 and were approached for participating in ALSPAC. Recruitment was conducted through media campaigns, visits to community centres, and distributing flyers at antenatal and maternity care centres. Out of 20 248 eligible pregnancies, 14,541 mothers initially enrolled in the first recruitment phase in between 1991-1992, and 452 eligible mothers and children retrospectively enrolled in the second recruitment phase, after re-approaching eligible women 7-years post-pregnancy (Boyd et al., 2013). Overall, 15,247 mothers and their children from phase 1 and phase 2 enrolled in ALSPAC.

2.6.1.3 Genotype data

Genotype data from ALSPAC was used in Chapters 3 and 4. Maternal blood during pregnancy was taken for maternal genotyping and offspring genotyping was derived from cord blood and peripheral blood taken at the age of 7 (Pembrey, 2004). As of February 2019, the ALSPAC DNA bank contains DNA of 7,100

offspring, 6,000 mothers and 2,400 partners (http://www.bristol.ac.uk/alspac/researchers/our-data/biological-resources/). The ALSPAC children genotype data has been generated using the Illumina HumanHap550 quad chip genotyping platforms by 23andme subcontracting the Wellcome Trust Sanger Institute, Cambridge, UK and the Laboratory Corporation of America, Burlington, NC, US. Mothers were genotyped using Illumina human660w quad array at the Centre National Genotypage and genotypes were called with Illumina GenomeStudio. Quality control filtering was done with the PLINK (v1.07) software. SNPs with a minor allele frequency of < 1%, call rate < 95% and Hardy-Weinberg equilibrium $P < 5 \times 10^{-7}$ were removed. Both, offspring and maternal genotype data have been jointly imputed to the 1,000 genomes reference panel (version 1, Phase 3, Dec 2013 Release).

2.6.1.4 DNA methylation data

DNA methylation data from ALSPAC was used for analyses in Chapters 4 and 5. The ALSPAC sub-sample that has been selected for DNA methylation assessment is called ARIES. A detailed description of the sample can be found elsewhere (Relton, Gaunt, et al., 2015). Briefly, DNA methylation data has been collected for 1018 mother-child pairs, based on maternal and offspring availability of genotype data. Specifically, maternal genotype data needed to be available at offspring's birth (collected at antenatal clinic) and at offspring's mean age of 15.5 years. Offspring's genotype data needed to be available at the same time-points (birth and 15.5 years), as well as at mean age 7.5 years. For offspring, DNA methylation at birth was extracted from cord blood and from peripheral blood at mean age of 7 and 15.5 years. In ARIES, genome-wide DNA methylation has been assessed through the Illumina Infinium HumanMethylation450 BeadChip (450 K) array. ARIES has been normalised using functional normalisation and estimated cell proportion types were generated using the Houseman method, which both was implemented through the R-package Meffil (Min et al., 2018). In this thesis, cord blood and peripheral blood DNA methylation at age 7 were used.

2.6.1.5 Phenotype data

Phenotype data of ALSPAC was used in Chapter 3, 4 and 5. Phenotype data was collected through questionnaires, which can be divided into "Carer

questionnaires", that were answered by the carer of the child (mostly mothers, 22 questionnaires to date), "Child based questionnaires", that asked questions about the child but were completed by carers of the child (mostly mothers, 26 questionnaires to date), "Child completed questionnaires", that children completed answering questions about themselves (30 questionnaires to date), "Partner questionnaires", that the partner of the carer of the child filled in about themselves (17 questionnaires to date), "Father questionnaires", that fathers of the child completed after enrolling themselves into the study (one questionnaire to date), "Puberty questionnaires", that were initially filled in by the carer of the child then by the carer and child together and finally by the child alone (nine questionnaires to date), and "School questionnaires", that were completed by school staff about the child (13 questionnaires to date). See http://www.bristol.ac.uk/alspac/researchers/our-data/questionnaires/ for an overview of the questionnaires.

In addition to questionnaires, 10 focus clinics for offspring were conducted, where biological samples and clinical assessments were collected. Mothers were invited to 4 focus clinics and fathers were invited to one focus clinic (see http://www.bristol.ac.uk/alspac/researchers/our-data/clinical-measures/). The phenotype data that was used for the analyses of this thesis will be discussed in more detail in the corresponding chapters.

2.6.1.5.1 Maternal smoking and caffeine consumption during pregnancy

In ALSPAC, mothers self-reported their smoking behaviour during pregnancy by filling-in questionnaires sent out at 8-, 18-, and 32-weeks of gestation, as well as at 8-weeks post-pregnancy. In Chapter 3, the following smoking during pregnancy variables were used: Ever smoked during pregnancy, stopped smoking during pregnancy, or cut down smoking during pregnancy (binary items assessed at 8-weeks gestation), and ever smoked during the first three months of pregnancy (binary assessment at 18 weeks gestation). For the analysis of Chapter 4 and 5, where smoking during pregnancy was used as a covariate, smoking was coded as a categorical variable representing 0 = no or early smoking during pregnancy, 1 = stopped before the second trimester of pregnancy and 2 = smoking in the third trimester or throughout pregnancy. The categorical smoking during pregnancy

variable was based on questions about smoking during pregnancy asked in questionnaires at 18-weeks gestation ("*Did you smoke regularly during the first three months of pregnancy*" and "*have you smoked during the past two weeks*"), 32-weeks gestation ("*How many cigarettes per day are you smoking at the moment*?"), and 8-weeks post-pregnancy ("*How many cigarettes (pipes or cigars) per day did you smoke during the last two months of pregnancy*").

Maternal caffeine consumption in ALSPAC has been assessed through self-report in questionnaires at 8-, 18- and 32-weeks of gestation. Participants were asked how many cups of coffee and tea, and how many cans of cola they drank per week. In this thesis, I use the caffeine variables assessed at 18- and 32-weeks gestation to derive a variable capturing consumption of caffeine in milligrams per day (mg/day). For the analyses of Chapter 3, variables assessing changes in caffeine consumption since finding out about the pregnancy have also been included (binary assessment at 8-weeks gestation). For each caffeinated drink (coffee/tea/cola), participants were asked whether they stopped, reduced, craved/had more, or never had drunken the caffeinated drink during pregnancy.

2.6.1.5.2 Offspring mental health outcomes

This thesis used ALSPAC data of offspring phenotypes assessed in childhood (3 to 10 years of age) and adolescence (12 to 18 years of age). The mental health phenotypes in childhood that were used in Chapters 3 and 5 were assessed through questionnaires based on maternal report. Internalising and externalising problems were assessed through the Strengths and Difficulties questionnaire (SDQ) (Goodman, 1997), the Development and Well-being Assessment (DAWBA) (Goodman et al., 2000), and the Short Mood and Feelings questionnaire (SMFQ) (Angold et al., 1995). Internalising phenotypes included: Symptoms of anxiety disorder (DAWBA), emotional problems (SDQ), specific phobias (DAWBA), depression (SMFQ), and post-traumatic stress disorder (DAWBA). Externalising phenotypes included: ADHD (SDQ), conduct disorder (SDQ), and oppositional defiant disorder (DAWBA). The adolescence phenotypes used in Chapter 3 were based on self-report, in addition to maternal report. Self-reported symptoms of depression, anxiety, and phobia were assessed using the revised Computerised Interview Schedule (CIS-R) (Lewis et al., 1992) during clinic visits at the age of 17. Self-reported psychosis symptoms (positive and

negative symptoms) were assessed using the psychosis-like symptoms (PLIKS) semi-structured interview (Horwood et al., 2008) during focus clinics at the age of 13 and 17. Diagnosis of eating disorders were assessed through adolescents self-reporting on an item that asked whether they have been treated for an eating disorder by a doctor, nurse, or health care provider at the age of 13 and 16. Self-harm was self-reported through multiple items asked in questionnaires at the age of 16 (see Easey, Mars et al., 2019 for more information).

2.6.2 Generation Rotterdam (Generation R)

2.6.2.1 Overview

The Generation R Study is a prospective birth cohort study that recruited pregnant women with an expected delivery date between 2002-2006 that were living in Rotterdam the Netherlands. A more detailed description of the cohort is provided in Kooijman (Kooijman et al., 2016) and Kruithof and colleagues (Kruithof et al., 2014). The motivation to set up this cohort was to understand the contribution of environmental and genetic factors to physical and mental health outcomes to improve prenatal and postnatal healthcare. The main outcomes of interest are categorised into growth and physical development, behavioural and cognitive function, childhood illnesses, and health and health care use (Hofman et al., 2004). The aim of the cohort was to recruit data of 10,000 offspring and to include other ethnicities in addition to Dutch ethnicities, such as other European, Moroccan, Surinamese, and Turkish ethnicities (Hofman et al., 2004). The Medical Ethical Committee of Erasmus MC, University Medical Center Rotterdam gave ethical approval for the data collection of this study.

2.6.2.2 Participants and data collection

Recruitment was conducted through midwifes and obstetricians that distributed information about the study (orally and through handouts) to eligible pregnant women who were later on provided with additional information by the staff of Generation R either over the phone or a home visit, as well as during their first routine ultrasound appointment (Hofman et al., 2004). For recruitment of non-Dutch speaking pregnant women, information leaflets were translated to other languages and staff was recruited that could, besides Dutch, also communicate in English, French, Portuguese or Turkish. At baseline, the cohort included 9,778

pregnant women (response rate 61%), which are still followed-up (response rate 10 years post-pregnancy = 80%) (Kooijman et al., 2016). The baseline sample included 39% of participants identifying as Dutch, 13% as Turkish, 11 % as Moroccan, 10% as Surinamese, 4% as Anilles, 3% as Cape Verde, and 20% as Others (Hofman et al., 2004).

2.6.2.3 DNA methylation assessment

Offspring was selected for DNA methylation assessment based on European ancestry as measured by genetic principal components and oversampled from a subsample of children for whom more in-depth measurements were available at the age of 9 (Focus 9). Offspring DNA methylation was assessed at three time-points, including birth (derived from cord blood), age 5 and age 9 (derived from peripheral blood) using the 450k array. Data preparation and normalisation was conducted in R, following the CPACOR workflow (Lehne et al., 2015). After running quality control checks for technical variation, 469,242 CpG sites of 1,396 participants at birth, 493 at age, 5 and 465 at age 9 remained for analysis. Cell type proportions were estimated through the R-package minfi (Aryee et al., 2014), using the Houseman method (Houseman et al., 2014). In this thesis DNA methylation data collected at birth and peripheral blood in childhood (age 5) has been used in Chapters 4 and 5.

2.6.2.4 Phenotype data

A general overview of the phenotype data collected in the Generation R study can be found at https://Generation R.nl/researchers/data-collection/. Planned assessment phases of Generation R can be summarised into (1) a prenatal assessment phase, (2) assessment age 1-4 years, (3) assessment age 4-12 years and (4) assessment age 12-20 years. Recruitment is still on going, with the most recent assessment time point being assessments of offspring at the age of 17 years. In addition to physical and biological assessments of parents, mothers received four and fathers received one questionnaire via post during pregnancy. The study was initially set out with a main focus on understanding factors impacting offspring's growth development. Thus, in the initial assessment during pregnancy, in week 12, 20, and 30 of gestation, mothers were asked about fetal growth and related factors (Hofman et al., 2004). Between 0-4 years postnatally mothers received

eight and fathers received one questionnaire(s) via post. The following assessments were conducted during focus clinics for which participants visited one of the Generation R research facilities in Rotterdam. Focus clinics took place at age 5, 9, 13, and 17. In this thesis, Generation R data that was used included maternal phenotypes during pregnancy and offspring phenotypes at the age of 0-4 and focus 5.

2.6.2.4.1 Maternal caffeine consumption during pregnancy

In Generation R, maternal caffeine consumption from coffee and tea was assessed through self-report in questionnaires at week 18-25 of gestation and used in the analyses of Chapter 4. Participants were asked whether they consumed coffee and/or tea. If they answered yes, follow-up questions about whether the drink was caffeinated/decaffeinated and how many cups they consumed were asked (see Appendix K, K4 for more information).

2.6.2.4.2 Offspring mental health outcomes

Generation R contributed information about childhood internalising problems for the analyses of Chapter 5. Mothers reported on their offspring's emotional and behavioural problems at the age of 36 and 72 months by filling-in questionnaires. The internalising subscale of the preschool Childhood Behaviour Checklist (*CBCL-1½-5*; Achenbach & Rescorla, 2000) was used to assess internalising problems.

2.6.3 Norwegian Mother and Child Cohort Study (MoBa)

2.6.3.1 Overview

MoBa is a nation-wide birth cohort that aimed at recruiting as many pregnant women as possible that are living in Norway. The Norwegian Institute of Public Health manages the cohort and details about the study can be found at https://www.fhi.no/en/studies/moba/. Motivated by the fetal origins of health and disease hypothesis, the MoBa cohort was established in the late 90s, with the aim of extending available data to inspect a larger variety of potential early life exposures that may be relevant for offspring's development of diseases (Magnus et al., 2016). While the initial cohort that started in 1999 was only aimed at investigating offspring outcomes up until the age of 3-years, follow-up of participants has later on been extended to assess outcomes throughout

adolescence. Also, the focus has shifted from mainly studying health outcomes in offspring to also collecting more data of their parents (Magnus et al., 2016). The aim to collect data on 100,000 participants has been reached in 2009 and follow-up is still ongoing (Magnus et al., 2016). The Norwegian Data Inspectorate and The Regional Committee for Medical Research Ethics gave approval for MoBa data collection. Ethics approval was obtained from the Regional Committee for Ethics in Medical Research, Norway. The Institutional Review Board of the National Institute of Environmental Health Sciences, USA, further approved collection of MoBa1 and MoBa2 data.

2.6.3.2 Participants and data collection

Recruitment was conducted at 50 out of 52 Norwegian hospitals that contain maternity units between 1999 and 2009. Information about the study, as well as the consent form and baseline questionnaire, were sent out along with invitations for routine ultrasound appointments (Magnus et al., 2006). Participation was restricted to Norwegian speaking participants as all the information was only provided in Norwegian only. Data of more than 114,000 offspring, 95,000 mothers and 75,000 fathers has been collected (Magnus et al., 2016).

2.6.3.3 DNA methylation assessment

A random subset of offspring was selected for DNA methylation assessment, which is referred to as MoBa1 and recruited DNA methylation samples between 2002-2004 (Joubert et al., 2016). In addition to a random subset, two non-random subsets of the MoBa cohort, called MoBa2 and MoBa3, were selected for DNA methylation assessment. MoBa2 was designed to study associations with childhood asthma and thus included cases and controls for asthma that were recruited between 2000-2005. MoBa3 aimed to study associations between DNA methylation and childhood cancer, which were recruited between 2000-2008. The analyses of this thesis only entail MoBa1 and MoBa2 data that have been used in Chapter 4 and 5. In both samples, DNA methylation was assessed using the 450k microarray. Beta values were calculated using the minfi R-package (Aryee et al., 2014) and data was normalised using the intra-array normalization strategy Beta Mixture Quantile dilation (Teschendorff, Marabita, et al., 2013). After quality control checks, Moba1 contained 1,068 children and MoBa2 685 children.

2.6.3.4 Phenotype data

In each trimester of pregnancy mothers were asked to fill in questionnaires and fathers were asked to fill in information once during their partner's pregnancy. Postnatal assessments included questionnaires at 6 and 18 months, and at 3, 5, 7, 8, 13, and 14 years post-pregnancy (see https://www.fhi.no/en/studies/moba/for-forskere-artikler/questionnaires-from-moba/ for a detailed overview of each questionnaire). In 2015 and 2016, an additional questionnaire was sent out to fathers asking a variety of questions about the father's physical and mental health as well as the relationship to their child and the child's biological mother. Biological samples were collected at the routine ultrasound appointment. Additionally, several national health registries have been linked to MoBa participants through their national identification number (Magnus et al., 2016).

2.6.3.4.1 Maternal caffeine consumption during pregnancy

The analyses of Chapter 4 included variables of maternal caffeine consumption during pregnancy assessed in MoBa1 at 17-weeks of gestation. In MoBa, coffee consumption was assessed by asking participants how many cups of instant coffee and/or espresso they consumed and what type of brewing method they used (boiled/percolated/filtered). Caffeinated tea consumption was also assessed in cups per day. Cola/pepsi consumption was assessed in mugs per day (see Appendix K, K3 for more information).

2.6.3.4.2 Offspring mental health outcomes

In Chapter 5, information on childhood internalising problems assessed in MoBa1 and MoBa2 was used. In a questionnaire sent out at 36-months post-pregnancy, mothers were asked to report on their offspring's emotional and behavioural problems. Questions were based on nine selected items of the internalising subscale of the preschool Childhood Behaviour Checklist (CBCL-1½-5) (Achenbach & Rescorla, 2000; see Appendix T, T3 for more information).

2.6.4 Born in Bradford (BiB)

2.6.4.1 Overview

BiB is a prospective birth cohort study that was initiated in 2007 and recruited 12,453 pregnant women residing in the city of Bradford, UK. Mothers were invited to participate in BiB during a routine oral glucose tolerance test (OGTT)

in their second trimester of pregnancy (26-28 weeks gestation). More detailed information about the cohort can be found in the paper by Wright and colleagues (Wright et al., 2013) or on the studies website (https://borninbradford.nhs.uk/research/). Bradford is one of the most disadvantaged areas in the UK and the cohort was established with the aim to tackle and investigate causes for increased childhood mortality in the area, compared to the rest of the UK (Wright et al., 2013).

2.6.4.2 Participants and data collection

Approximately 20% of the population in Bradford are of South Asian ethnicity, predominately stemming from Pakistan (90%). To be able to include participants that could not speak English, the baseline questionnaires were translated into Urdu and Mirpur, and translators were provided for facilitating completion of the questionnaires. Out of 12,453 mothers that were recruited, 13,776 had successful pregnancies, of which 11,396 (82.7%) of mothers completed the baseline questionnaire. In addition to maternal data, data of 3,448 fathers were collected at baseline. Two subgroups of BiB were followed-up after pregnancy, one focussing on risk factors for childhood obesity in children born between October 2008 and May 2009 (N = 1,700) and one focussing on viral infections and allergic diseases (N = 2,300). The Bradford Research Ethics Committee (Ref 07/H1302/112) granted ethics approval for the data collection.

2.6.4.3 DNA methylation data

DNA methylation data of BiB was used in Chapter 4. Out of the participants with a singleton birth that (1) had completed the OGTT (2) had maternal and offspring genotype data available, and (3) self-identified as Pakistani or white British ethnicity (together representing ~90% of sample), a subsample of 500 Pakistani and 500 white British mother-child pairs were randomly selected for DNA methylation assessment (total N = 2,000). DNA methylation was assessed using the EPIC array derived from maternal blood samples taken at recruitment and offspring's cord blood samples. After quality control checks for technical variation and matching genotype probes with genotype data, 864 samples remained. Data was normalised using functional normalisation and cell proportion types were estimated using the Houseman method, both conducted within the R-

package meffil (Min et al., 2018). Cord blood DNA methylation data of BiB has been used for the analysis in Chapter 4.

2.6.4.4 Phenotype data

Phenotype data of BiB that was used in Chapter 4 of this thesis and only included maternal phenotypes assessed during pregnancy. A general overview of the phenotype data available in BiB can be found at https://borninbradford.nhs.uk/research/documents-data/. Briefly, baseline questionnaire data is available for 10,519 mothers and 3,287 fathers. Offspring phenotype data is available for the subsamples that were selected for follow-up (see BiB overview). Additionally, 99.1% of mothers and 90% of offspring participants were successfully linked to Bradford Royal Infirmary hospital records, giving access to a wide range of maternity and birth related phenotypes. Also, 84% of offspring participants could be linked to education records.

2.6.4.4.1 Maternal caffeine consumption during pregnancy

Data of maternal caffeine consumption assessed in BiB was used in the analyses of Chapter 4. In BiB, participants self-reported on the number of cups of caffeinated tea, filter/cafetiere coffee, and instant coffee they consumed per day at 26- to 28-weeks of gestation (see Appendix K, K2 for more information).

2.6.5 INfancia y Medio Ambiente (Environment and Childhood; INMA)

2.6.5.1 Overview

The aim of the INMA cohort was to investigate prenatal and postnatal factors that influence offspring growth and health development. A more detailed description is provided in the cohort profile (Guxens et al., 2012). The cohort includes data collected in several areas across Spain (Ribera d'Ebre, Menorca, Granada, Valencia, Sabadell, Asturias and Gipuzkoa), according to a common data collection protocol and is managed by the Barcelona Institute for Global Health (for more information also see https://www.proyectoinma.org/en/inma-project/).

2.6.5.2 Participants and data collection

Participation was restricted to mothers over the age of 16, who conceived without assisted reproductive technologies, did not experience any communication problems, lived in one of the recruitment areas, and were expected to be

delivering a singleton birth in the hospital where recruitment took place (Guxens et al., 2012). The recruitment time-point varied by study location and took place between 1997 and 2006 and the response rate ranged by study location from 45% to 98%. Furthermore, women recruited in the different areas differed on various demographic characteristics (age, education, employment status) (Guxens et al., 2012). Data was collected in the first and second trimester of pregnancy and at 6-assessment time points between birth and offspring at the age of 10 years (assessment time-points: 6-months; 1-1.5, 2-2.5, 4-5, 6-7, 9-10 years). Questionnaire data and sample size varied between included studies in INMA.

2.6.5.3 Maternal caffeine consumption during pregnancy

Data of maternal caffeine consumption assessed in INMA was used in the analyses of Chapter 4. Questionnaires were sent to participants to assess maternal coffee, tea, and cola consumption at 12-weeks of gestation. A categorical variable, ranging from "*never or less than once per month*" to "*six or more per day*" was used to assess the average amount of cups participants consumed (see Appendix K, K5 for more information).

2.6.5.4 DNA methylation assessment

In INMA, DNA methylation data was available for 385 out of the 742 children that enrolled in one of INMA's sub cohorts (called Sabadell). Out of the families that consented to collection of cord blood for genetic and epigenetic assessments, a random subsample of 521 children was selected. The final sample comprised of 385 children that were selected based on good DNA quality, availability of genetic and other data, and European ancestry. Offspring DNA methylation was assessed at birth (cord blood) and at the age of 4 years (peripheral blood) using the 450k microarray. Beta-values were generated through Illumina's analysis software GenomeStudio and data was normalised using the minfi R-package (Aryee et al., 2014).

2.6.6 Etude des Déterminants pré et post natals du développement et de la santé de l'Enfant (EDEN)

2.6.6.1 Overview

Similarly to the other cohort studies, EDEN (study on the pre- and early postnatal determinants of child health and development) (Heude et al., 2016) was set up to

investigate the influence of pre- and postnatal exposures through the maternal environment on offspring's health development (Heude et al., 2016). EDEN started recruiting pregnant women in 2003 through two University Hospitals in France (Nancy and Poitiers) with a maternity unit. Mother-child pairs have been followed-up for up to 8-years post-pregnancy. Ethics approval for data collection was granted by the ethics committee of Kremlin Bicêtre and the Commission Nationale Informatique et Liberté.

2.6.6.2 Participants and data collection

Women were accounted as eligible for the study if they attended the maternity unit before week 24 of amenorrhoea. Further eligibility criteria were a singleton pregnancy, no diagnosed diabetes before pregnancy, French literacy and not planning to move outside of the recruitment area within the next 3 years. Out of 3,758 eligible women that were recruited between 2003 and 2006, 1,034 women enrolled in Nancy and 968 in Poitiers, resulting in a total sample size of 2,002 pregnant mothers (response rate = 53%). Data about maternal and offspring data was collected at 3 clinic visits for mothers, (time-points: 24-28 weeks of amenorrhea, delivery, 5-6 years post-pregnancy), and 4 clinic visits for offspring (time-points: birth, 1-, 3-, 6-years of age). Furthermore, data was collected through one prenatal questionnaire (24-28 week amenorrhoea) and 8 postnatal questionnaires (time-points: 4-months; 8-months; 1-, 2-, 3-, 4-, 5- and 8-years postnatally).

2.6.6.3 Maternal caffeine consumption during pregnancy

Phenotype data of Eden was used in the analyses of Chapter 4. In weeks 24-28 of gestation questionnaires were sent out to participants to assess their consumption of coffee, tea, and cola. Participants were asked about the average amount of daily cups consumed during the first trimester of pregnancy (see Appendix K, K6 for more information).

2.6.6.4 DNA methylation assessment

For the DNA methylation subsample a random sample of 150 children was selected amongst children who met the following requirements: (1) Participated in the five-year follow-up assessment, (2) had consent for the collection of cord blood and peripheral blood at the age of five, and (3) had a DNA methylation

sample with sufficient quality. Data was assessed using the 450k microarray and after quality control checks, 439,306 CpG sites were available for analysis (Merid et al., 2020). Estimated cell proportion types were estimated using the R-package minfi (Aryee et al., 2014). PCA was conducted in EDEN and 5 PC were used to control for batch effects (Merid et al., 2020).

2.7 Databases

For contextualising the results obtained from my analyses I consulted publicly available data resource platforms, which are introduced below. Each of these platforms provides a user-friendly overview over the publicly available scientific evidence for (epi)genetic associations with a variety of traits.

2.7.1 Genecards

The Genecards database (https://www.genecards.org/) provides an overview over the information that is available for all annotated genes. The database is managed by the Department of Molecular Genetics at the Weizmann Institute of Science in Israel. Genecards pools data from over 150 web-based platforms with information about genomic, transcriptomic, proteomic, genetic, clinical and functional information (see https://www.genecards.org/Guide/Sources for an overview of the genecards sources). I used the genecards database to understand the biological function of the genes annotated to the differentially methylated CpG sites found in my analyses and to explore whether there might be a potential biological pathway from changes in the expression of these genes to different manifestations in my traits of interest.

2.7.2 NHGRI-EBI GWAS Catalog

The NHGRI-EBI GWAS catalogue, or short GWAS catalog, can be accessed through https://www.ebi.ac.uk/gwas/ (Buniello et al., 2019). The catalogue provides an overview over all published GWAS that fulfil a sufficient scientific standard to be added to the catalogue (e.g., genome-wide analysis instead of candidate gene analysis, see https://www.ebi.ac.uk/gwas/docs/methods/criteria for all eligibility criteria). Studies are identified by weekly literature screenings, curated and added to the GWAS catalog by expert scientists (see https://www.ebi.ac.uk/gwas/docs/methods/curation for more detail on the curation procedure). The database was established in 2008 and over 10 years later, in 2019,

contained over 70,000 SNP-trait associations (Buniello et al., 2019). The catalogue's website contains a search engine were GWAS results of specific traits, variants, genes or genomic locations of interest can be looked-up. The corresponding GWAS study summary statistics can also be downloaded for further analysis. Evidence for an SNP-trait association is determined by a P-value of less P-value < 1.0×10^{-5}. In my thesis, I used the GWAS catalog to check whether my traits of interest showed associations with SNPs annotated to the same genes, as the genes that were annotated to differentially methylated CpG sites in my EWAS analyses. As more GWAS studies than EWAS studies have been conducted and DNA methylation is partly influenced by genetics, any congruent evidence between genes detected in GWAS and EWAS studies could provide further evidence for a biological function of that gene on the traits of interest.

2.7.3 EWAS catalog

The EWAS catalog was established by the MRC Integrative Epidemiology Unit at the University of Bristol (http://www.ewascatalog.org/). Similar to the GWAS catalog, the EWAS catalog is a platform that pools summary statistics from all published EWAS studies of DNA methylation. Currently the catalogue is based on three main data sources, including: (1) Studies extracted from the literature through regular screenings of PubMed using the R-package *journalclub* (Suderman & Yousefi, 2020), (2) studies that have been conducted (but are not all published) within the Bristol-based ARIES DNA methylation sample, and (3) EWAS studies conducted within the Gene Expression Omnibus database (available at https://www-ncbi-nlm-nih-gov.bris.idm.oclc.org/geo/). Studies are classified as eligible for the catalogue if they were published after 2010, analysed at least 100,000 CpG sites, contained at least 100 participants, and include new data. Evidence for a CpG–trait association was determined by a P-value of less than 1×10^{-4} (http://www.ewascatalog.org/documentation/). In this thesis, I am using the EWAS catalog to explore (1) if the differentially methylated CpG sites of my analyses show associations with any other traits that might be related to my trait of interest, (2) whether the genes annotated to the differentially methylated CpG sites found in my analyses are associated with any other traits, and (3) identify previously published EWAS studies of my traits of interest.

2.8 Chapter summary

In this chapter, I have provided an introduction to the methods and samples used in this thesis. The methods that were introduced are MR and PRS analysis, as well as EWAS. MR is a useful technique that uses genetic instruments as proxies for exposures to enable causal inferences from observational data (Davey Smith & Ebrahim, 2003). Using a PRS instead of single variants may help to increase predictive power of genetic instruments for complex traits by aggregating the small effects of many SNPs into one combined score (Choi et al., 2020). In contrast to investigating DNA methylation at specific candidate genes, EWAS is a hypothesis generating approach, which is useful for discovering novel differentially methylated CpG sites, as well as validating CpG sites previously found in candidate-gene studies (Jones et al., 2018). In addition to discussing the methods, an overview over the six birth cohorts to which these methods are applied has been given. The specific measures of smoking and caffeine consumption, as well mental health phenotypes used of each of these cohorts, are discussed in more detail in each corresponding results chapter. Lastly, the Genecards, GWAS and EWAS catalog have been introduced, which were consulted across this thesis for contextualising results with findings of previously published (epi)genetic studies.

Chapter 3 – Maternal and child genetic liability for smoking and caffeine consumption and child mental health: An intergenerational polygenic risk score analysis in the ALSPAC cohort

This chapter includes sections of the following manuscript that has been published on medRxiv and the Addiction Journal:

Schellhas, L.*, Haan, E.*, Easey, K. E., Wootton, R. E., Sallis, H. M., Sharp, G. C., Munafò, M. R., & Zuccolo, L. (2021). Maternal and child genetic liability for smoking and caffeine consumption and child mental health: An intergenerational genetic risk score analysis in the ALSPAC cohort. *Addiction*. https://doi.org/10.1111/add.15521

* contributed equally to this work

This study was part of a larger collaborative project conducted as a joint effort between PhD students and post-doctoral researchers, which included analyses of alcohol, caffeine, and smoking polygenic risk scores (PRS). Phenotype extraction and preparation of the variables for the analysis were divided between Dr. Kayleigh Easey (PhD student at the time), Dr. Robyn Wootton (post-doctoral researcher), Dr. Elis Haan (PhD student at the time), and myself. I extracted and prepared the caffeine variables, as well as IQ, personality, schizophrenia, sleep, physical activity, and body perception phenotypes for mothers during pregnancy and offspring. The script for generating PRS was provided by Dr. Robyn Wootton. Smoking PRS (for mothers and offspring) were generated by Dr. Robyn Wootton, the caffeine PRS for mothers was generated by Dr. Elis Haan, and the caffeine PRS for offspring was generated by myself. Individual datasets of each researcher were combined into a final dataset that was used for the analysis. Dr. Kayleigh Easey has written a general script for running the main analysis and conducted the analysis for the alcohol PRS, which has been published as a separate paper (Easey et al., 2021). For the main analysis of this project, I modified the script of Dr. Kayleigh Easey and conducted the analysis of the caffeine PRS and the lifetime smoking PRS. Dr. Elis Haan conducted the smoking initiation analysis. I have generated all the plots and figures presented in this

chapter. The tables have mostly been generated by Dr. Elis Haan, except for the results tables of the caffeine PRS (maternal and offspring) and the lifetime smoking PRS (offspring) that I have generated. I have drafted the introduction, results, and discussion section of this paper. Throughout the project, we double-checked and provided feedback on each other's work. Co-authors provided comments that were incorporated before submitting the paper for publication.

3.1 Chapter overview

As mentioned in Chapter 1, PRS analyses may be applied for investigating causal effects in observational study designs if the PRS is a valid proxy for the exposures of interest (here: caffeine and smoking). This is a specifically useful method for investigating causal effects of exposures where it is difficult to randomly assign participants to being exposed or unexposed, making it an attractive potential method for investigating effects of intrauterine exposures. However, there are several caveats when using MR to investigate intrauterine effects that are discussed and explored in more detail in this chapter. These include: lack of validation of the genetic variants as proxies for behaviours during pregnancy, genetic proxies lacking specificity for an exposure time point (they reflect lifetime maternal exposure instead of only exposure during pregnancy), and the potential of confounding due to the shared genetic liability between mother and offspring. The aim of this chapter was first, to test whether the PRS for smoking and caffeine consumption can be used as proxies for these behaviours during pregnancy and second, to explore, in a hypothesis free design, which mental health outcomes show evidence for causal (or confounded) associations with the smoking and caffeine PRS in ALSPAC (examining the exposure-disease relationship (step 1) of the meet-in-the-middle approach.) The second aim was explored by applying a cross-generational approach to a phenome-wide association study (PheWAS) design, which allowed to compare three effects: (1) offspring smoking and caffeine PRS on their own mental health outcomes, (2) maternal smoking and caffeine PRS on their own mental health outcomes, and (3) an intergenerational analysis testing the effect of maternal PRS on offspring mental health outcomes in childhood. The comparison of these three effects could be used to provide indications for: A potential causal effect of smoking and caffeine consumption on own mental health, pleiotropic associations of the smoking and caffeine genetic variants with mental health outcomes, or an effect of maternal smoking or caffeine consumption (intrauterine or postnatal) on to offspring.

3.2 Introduction

Smoking and caffeine consumption often co-occur (Treur et al., 2016) and are associated with mental health problems and other substance use behaviours

(Kendler et al., 2008; Lara, 2010). There is some evidence that smoking is a causal risk factor for mental health problems such as depression and schizophrenia (Vermeulen et al., 2019; Wootton et al., 2019); however, the relationship between caffeine and mental health is less clear, and possibly difficult to disentangle from smoking as the two often co-occur (Lara, 2010; Temple et al., 2017). In addition to associations between smoking, caffeine, and mental health outcomes within individuals, observational research suggests that prenatal maternal consumption of tobacco and caffeine could have an intergenerational effect on offspring's mental health (Bekkhus et al., 2010; Dolan et al., 2016; Moylan et al., 2015; Tiesler & Heinrich, 2014).

Using conventional epidemiological methods alone, it is difficult to ascertain whether prenatal tobacco and caffeine exposure causally affect offspring mental health outcomes (Gage et al., 2016; Thapar et al., 2009). Not only do mothers and offspring share a similar environment (such as socio-economic position), they also share, on average, 50% of their segregating genetic variation. Due to this shared genetic and environmental confounding it is difficult to disentangle the effect of maternal substance use on offspring mental health from those of offspring's own substance use.

The association between maternal prenatal smoking and internalising problems in children is less extensively researched compared to associations with externalising problems, and existing evidence is mixed (Ashford et al., 2008a; Menezes et al., 2013; Moylan et al., 2015; Taylor et al., 2017). Many studies report a positive association between prenatal smoking and offspring's externalising problems (Brion et al., 2010b; Dolan et al., 2016; D'Onofrio et al., 2008; Manzano et al., 2016), which could reflect a potential intrauterine effect of smoking. However, results vary when adopting different methods to account for shared environmental and genetic confounders (D'Onofrio et al., 2008). For example, studies using negative controls designs and sibling comparisons have found inconclusive evidence for a causal intrauterine effect (Brion et al., 2010; D'Onofrio et al., 2008; Langley et al., 2012; Nomura et al., 2010; Roza et al., 2009). In fact, study designs adjusting for shared genetic factors between mother and offspring have concluded that genetic factors explain associations between maternal prenatal smoking and externalising problems in offspring (Rice et al., 2018). This

literature highlights the complexity of the nature of associations between pregnancy exposures and offspring mental health, and the importance of disentangling shared genetic and environmental confounders to understand whether a true causal effect exists.

3.2.1 Intergenerational Mendelian Randomisation

Using polygenic risk scores (PRS) as proxies for smoking or caffeine consumption can, in principle, reduce bias from confounding (see Chapter 2) (Davey Smith et al., 2007). As with conventional MR, unbiased estimates from intergenerational PRS analyses, where maternal genetic variants are used to test associations with offspring outcomes, can only be derived under the assumption that the key MR assumptions hold. In the context of intrauterine smoking and caffeine exposure the assumptions are as follows (Diemer et al., 2020; Lawlor et al., 2017): First, the maternal smoking and caffeine PRS must be robustly associated with the behaviours during pregnancy (relevance assumption). Second, the maternal smoking and caffeine PRS are only associated with offspring mental health outcomes through the intrauterine exposure to smoking or caffeine (exclusion restriction assumption). Third, none of the confounding variables of the association between maternal smoking or caffeine consumption during pregnancy and offspring mental health outcomes may be associated with the maternal PRS for smoking or caffeine (exchangeability assumption). A visualisation of the application of intergenerational MR to research the effect of smoking during pregnancy can be found in Figure 3.1. There are several reasons why these assumptions could be violated in intergenerational MR (or PRS) analyses that investigate the effects of smoking and caffeine consumption during pregnancy on offspring mental health problems. First, the genetic variants used in the PRS have mostly been identified and validated in non-pregnant adult populations and thus might not predict behaviours during pregnancy (violating the first MR assumption) (Lawlor et al., 2017; Liu, Jiang, Wedow, Li, Vrieze, et al., 2019; The Coffee and Caffeine Genetics Consortium et al., 2015). This is the only assumption that can be tested directly by inspecting whether genetic variants that have been derived as proxies for behaviours in the general population, also predict the behaviour in pregnant populations (Lawlor et al., 2017). Second, offspring's own smoking or caffeine consumption may confound associations because

mothers may pass on their genetic predisposition for smoking or caffeine consumption to their children, or "environmentally" influence them in their behaviour (genetic nurturing/dynastic effects, Figure 3.1) (Kong et al., 2018; Pingault et al., 2018). Consequently, when offspring's mental health outcomes are assessed at an age where offspring are likely to have started smoking or drinking caffeine themselves, offspring's own consumption may cause offspring's mental health problems (violating the exclusion and exchangeability MR assumption because maternal PRS would not be independent of the confounding of offspring's smoking or caffeine use). Therefore, in order to avoid confounding of maternal exposure-offspring outcome associations through offspring's own smoking or caffeine consumption, it may be useful to apply a negative control approach that tests associations between maternal genetic variants and offspring outcomes before offspring start smoking or consuming caffeine themselves. Third, an association between maternal PRS and offspring mental health outcomes may reflect a shared genetic liability between smoking or caffeine consumption and mental health outcomes (pleiotropy, see Chapter 2) instead of a causal effect of the exposure (violating the exchangeability and exclusion restriction assumption because maternal PRS would not only be associated with offspring mental health through maternal smoking or caffeine use).

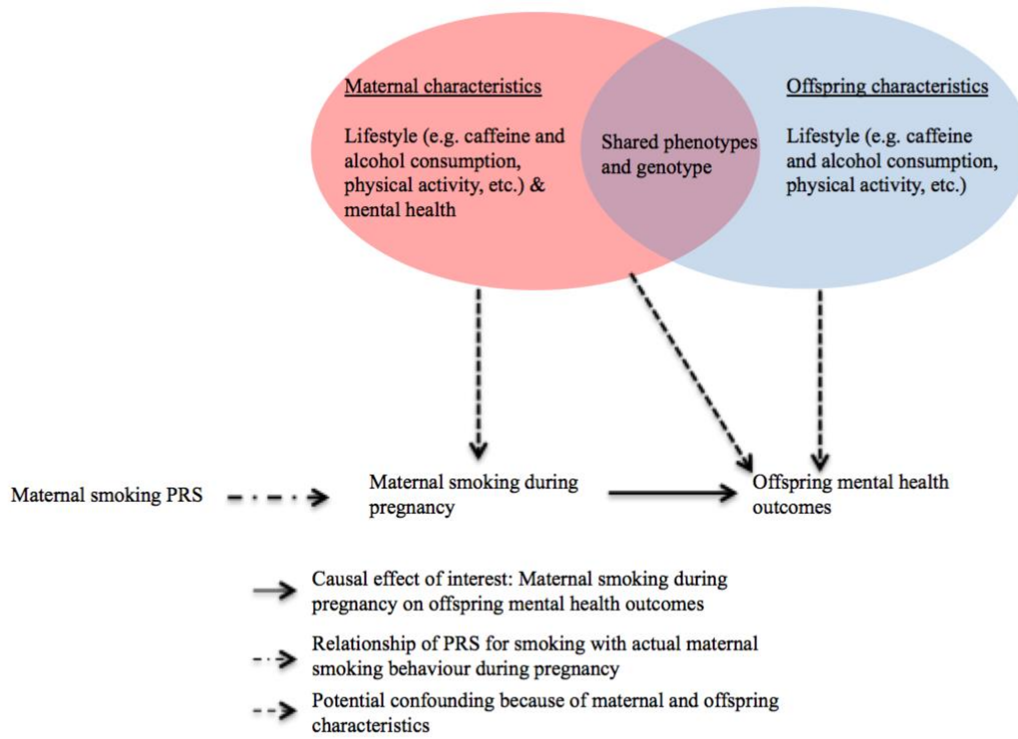


Figure 3.1 Application of intergenerational MR to investigate the intrauterine effect of smoking on mental health outcomes *(credit: adapted from Lawlor et al., 2017).*

3.2.1.1 Overlap between maternal and offspring genetic variants

Whereas pleiotropy is a well-known problem for MR analyses, the shared genetic predisposition between mother and offspring, that may confound associations between maternal genetic variants and offspring mental health outcomes, is a unique challenge for intergenerational MR analyses (Diemer et al., 2020). On the surface it may easily be accounted for by adjusting intergenerational PRS analyses for offspring PRS (Lawlor et al., 2017). However, this turns out to be a naïve approach, as the adjustment for offspring PRS may cause spurious association because it introduces collider bias through paternal genotype (Diemer et al., 2020; Lawlor et al., 2017). When two variables are causally related to a third variable, that third variable is called a collider (Munafò et al., 2018). When an analysis is adjusted (or stratified/selected) for a collider, bias may be introduced because a spurious association is created between the variables that cause the collider and the outcome (Griffith et al., 2020; Richmond et al., 2017). If an intergenerational MR analysis is adjusted for offspring genotype (the collider of maternal and paternal genotype, see Figure 3.2) an association might be created between maternal PRS and offspring outcome that is not attributable to the intrauterine exposure but the association between the unmeasured paternal PRS and offspring outcome (Lawlor et al., 2017).

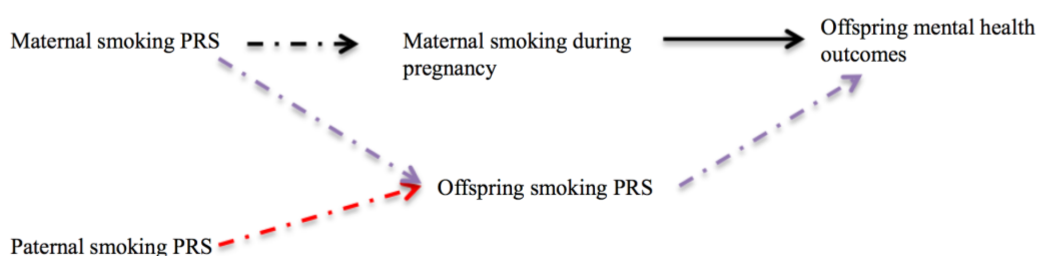


Figure 3.2 Illustration of offspring smoking PRS as a collider caused by maternal and paternal smoking PRS. *Spurious associations may be generated between maternal smoking PRS and offspring mental health outcomes if the analysis is adjusted for offspring smoking PRS.*

Therefore, the gold standard is that intergenerational PRS analyses should, in addition to offspring PRS, control for paternal PRS to properly account for the inherited genetic liability between mother and offspring. However, this was not possible in the analysis of this chapter due to lack of paternal genetic data available in ALSPAC to sufficiently power genetic analyses, which is a common problem in birth cohort studies (Lawlor et al., 2017). In order to avoid collider bias, but to still be able to separate maternal genetic effects from offspring genetic effects, maternal and offspring smoking and caffeine PRS were used as separate predictors for offspring mental health outcomes in this study.

3.2.2 Phenome-wide association studies (PheWAS)

The main method applied in this chapter is based on the concept of a phenome-wide associations study (PheWAS). PheWAS is a popular approach to discover novel and potential causal association between genetic variants (individually or as a PRS) and phenotypes in a hypothesis generating fashion (Bush et al., 2016; Millard et al., 2015). In contrast to GWAS, where associations between genetic variants and a single outcome are tested, a PheWAS tests the associations between preselected genetic variants (or a PRS) and a wide range of phenotypes (Bush et al., 2016). In addition to discovering new associations, it can be a useful approach for validating exposure-outcome associations discovered in GWAS, detecting pleiotropic effects, and to investigate the validity and specificity of a PRS (Bush et al., 2016; Hall et al., 2014). For instance, a PheWAS study that inspected associations between risk alleles for alcohol and nicotine use and a variety of outcomes in women (N = 26,394), was able to confirm already known genetic associations and discover novel associations, including psychological and socioeconomic traits (Polimanti et al., 2016).

The degree to which a PheWAS is hypothesis generating may vary depending on the specific research question and is determined by the number of phenotypes chosen. A comprehensive PheWAS tests the associations between genetic variants (or PRS) and the "entire" phenome, whereas a targeted PheWAS selects a certain group of phenotypes and tests these for associations with a genetic variant (or PRS) (Bush et al., 2016). The latter is more commonly applied for research questions with an underlying biological hypothesis that may influence a specific group of phenotypes (Bush et al., 2016). The PheWAS analysis of this chapter

follows a targeted PheWAS design, as the research question is specified to explore associations of the genetic variants of smoking and caffeine consumption specifically with mental health phenotypes. PheWAS is a powerful approach to complement more conventional hypothesis testing approaches that often are subject to publication bias, with only significant results being published, due to the fact that in a PheWAS all results are commonly reported, including null findings (Millard et al., 2015).

3.2.3 Chapter Aims

The analyses in this chapter aimed to elucidate the effects of maternal prenatal smoking and caffeine consumption on offspring mental health, using data from ALSPAC. The study had two specific aims: First, to validate the smoking and caffeine PRS during pregnancy (in mothers) and second, to estimate the effect of maternal smoking and caffeine consumption on offspring mental health outcomes. The second aim was achieved by first estimating the association between maternal smoking and caffeine PRS and offspring mental health outcomes during childhood (before age 10 years when children are unlikely to start smoking or consuming higher level of caffeine themselves; childhood PRS analysis, Figure 3.3), and then comparing the effect of mothers PRS and offspring PRS on offspring mental health. This comparison allowed disentangling pleiotropic from potential causal associations (intergenerational PRS analysis, Figure 3.3). Given the shared genetics between mothers and offspring, pleiotropic associations would be reflected by a larger estimated effect of the offspring PRS on offspring mental health, compared to the estimated effect of the maternal PRS (childhood PRS analysis, Figure 3.3). Following the same reasoning, a larger estimated effect of the maternal PRS on offspring mental health (relative to the estimated effect of the offspring PRS) would have provided more evidence to support a causal intrauterine effect of maternal behaviour on offspring mental health (intergenerational PRS analysis, Figure 3.3).

This study aimed at identifying offspring mental health phenotypes that show indication of being influenced by maternal smoking and caffeine consumption (pre-, during-, or post-pregnancy), instead of a shared genetic predisposition between mother and offspring. This chapter did not aim to specifically test a pre specified hypothesis that investigates the effect of maternal smoking or caffeine

consumption *during pregnancy* on a specific offspring mental health phenotype of interest. As genetic variants are allocated at conception and thus a genetic susceptibility for smoking and caffeine consumption influences lifetime consumption and not just consumption during pregnancy, the current analyses can neither differentiate between pregnancy effects, nor pre- or post-natal effects of smoking and caffeine consumption (Lawlor et al., 2017). Offspring mental health phenotypes that show evidence of an association with maternal smoking and caffeine consumption, which is unlikely to be explained by shared genetics between mother and offspring, should be followed-up to investigate specific timings of exposures (pre-, during-, or post-pregnancy). For instance, the specific influence of exposure to smoking or caffeine *during pregnancy* may then be followed-up by running subsequent MR analyses investigating a mediating role of offspring DNA methylation in cord blood (e.g., as I have done in Chapter 6).

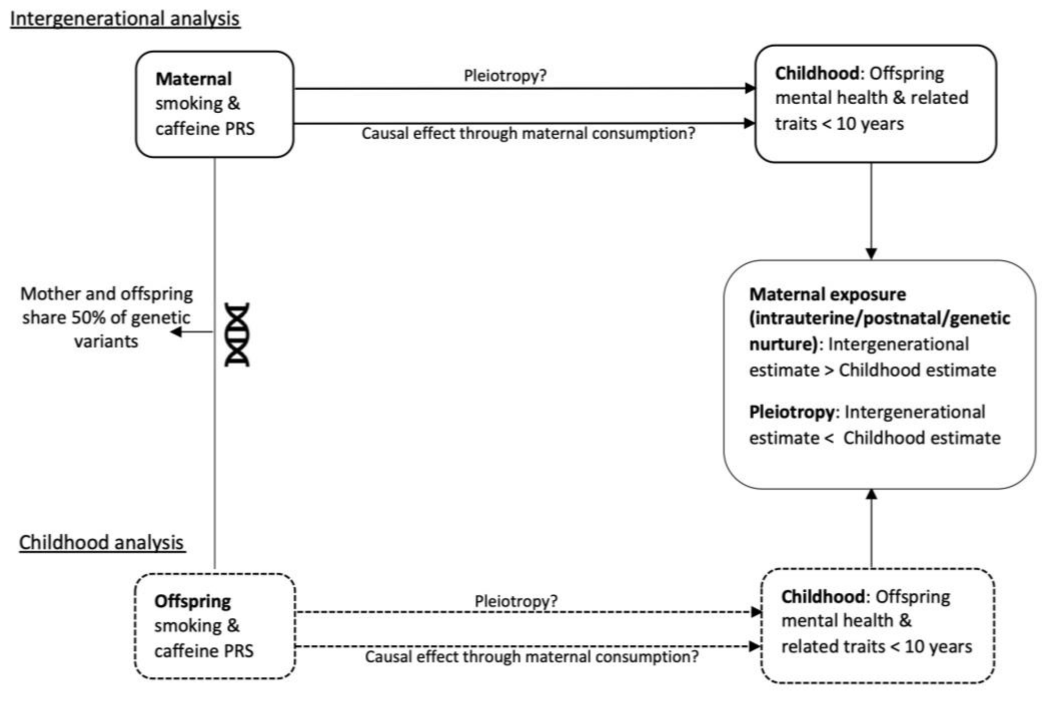


Figure 3.3 Design overview of the comparison of the intergenerational analysis (top) and the childhood analysis (bottom). *A larger effect estimate in the intergenerational compared to the childhood analysis would reflect a causal effect of caffeine/smoke exposure through the maternal environment. A larger effect estimate in the childhood compared to the intergenerational analysis would reflect pleiotropic association of the polygenic risk scores with mental health.*

3.3 Methods

3.3.1 Study population

ALSPAC data was used in the analyses of this chapter. For the cohort description please refer to Chapter 2.

3.3.2 Design

A visual overview of the study design can be found in Figure 3.3. Given the shared genetic material between mothers and offspring, evidence for pleiotropic associations should be reflected by a larger estimated effect of the offspring PRS on offspring mental health (childhood PRS analysis), compared to the estimated effect of the maternal PRS (intergenerational PRS analysis). Following the same reasoning, a larger estimated effect of the maternal PRS on offspring mental health (relative to the estimated effect of the offspring PRS) would provide more

evidence to support a causal effect of maternal behaviour on offspring mental health (intergenerational PRS analysis).

3.3.3 Phenotype data

Mental health phenotypes were selected from questionnaires and clinical assessments. Besides mental health phenotypes, some non-mental health phenotypes were also included, that were selected based on evidence in the literature indicating high comorbidity with mental health problems (e.g., alcohol, cannabis, other drugs, personality, body mass index, sleep, socio-economic variables). To validate the PRS, phenotypes describing caffeine consumption and smoking behaviours were derived. Offspring assessment points were grouped into 'childhood' (age 7-11 years) and 'adolescence' (age 12-18 years). Maternal assessment points were grouped into 'during pregnancy' (8, 18 and 32 weeks of gestation) and 'outside of pregnancy', which included phenotypes assessed pre- and/or post-pregnancy. Outcomes assessed within the first four years after pregnancy were excluded, as the transition to parenthood may influence mental health temporarily (Saxbe et al., 2018) and mothers may be more likely to be pregnant again. In total, 71 phenotypes for offspring (childhood and adolescence) and 79 phenotypes for mothers (during and outside of pregnancy) were included. Table 3.1 gives an overview of phenotypes included in the intergenerational and childhood PRS analyses across time-points. A complete list of phenotypes is given in Appendix A, Table A1.

Table 3.1 Availability of phenotypes included in the intergenerational and childhood analyses across the two generations

	Offspring		Mothers	
Measures	Childhood (age <10)	Adolescence (age 12-18)	Outside of pregnancy (pre/post-pregnancy)	During pregnancy
<u>Mental health</u>				
Emotional problems				
Depression symptoms	x	x	x	x
Anxiety symptoms	x	x	x	x
Specific phobia	x	x		
Behavioural problems				
ODD symptoms	x	x		
Conduct disorder symptoms	x	x	Personality measures (extraversion, anger, impulsivity)	
ADHD symptoms	x	x		
Total behavioural difficulties	x	x		
Neuro-developmental				
Autism (lifetime diagnosis)	x			
<u>Other</u>				
Handedness (negative control)	x			
IQ	x	x	Only education & SEP	Only education & SEP
Number of stressful life events	x	x	x	x
BMI	x	x	x	Only Image perception and physical activity
Sleep initiation	x	x		
Sleep maintenance	x	x		
Hours of sleep (duration)	x	x	x	
Overall caffeine intake	x	x	x	x

Note. ODD = Oppositional defiant disorder, ADHD = Attention deficit hyperactivity disorder, IQ = Intelligence quotient, BMI = Body mass index, SEP = Socio-economic position.

3.3.3.1 Phenotype selection and coding

3.3.3.1.1 Phenotype assessment

The phenotype selection was aimed at including as many different mental health phenotypes available in ALSPAC as possible. If phenotypes correlated highly, only the one with the larger sample size was included. Following the same reasoning, if some phenotypes were assessed at close time points (e.g., at age 14 and age 16), but only showed a low to moderate association, both time points were included. Continuous phenotypes that substantially deviated from a normal distribution were transformed into quantiles and if a validated cut-off score was available, binary phenotypes were derived. Zero inflation was accounted for by transforming continuous phenotypes with more than 20% of zero values into three quantiles (0, < median, > median). Caffeine phenotypes (daily consumption of coffee, tea, cola in milligrams per day (mg/day)) were generated based on their caffeine content. In ALSPAC, caffeinated coffee and tea were initially assessed in cups per day and caffeinated cola was assessed in cans per day. Caffeine content for each cup/can were transformed based on the assumption that one cup of tea contains 27 mg, a cup of coffee 57 mg, and a can of cola (330ml) 20 mg of caffeine (Farrow et al., 1998). Total caffeine consumption was computed by summing the caffeine content in mg/day of each drink. Extreme outliers, such as consuming more than 28 cups of coffee and tea per day, were removed.

3.3.4 Polygenic risk scores (PRS)

In ALSPAC, genome-wide SNP data were available for 8,237 children and 8,196 mothers. After removing individuals who withdrew their consent or did not pass quality control, PRS could be generated for 7,964 children and 7,921 mothers (see Chapter 2 for more information about genotyping in ALSPAC) (Taylor, Jones, et al., 2018). The GWAS and Sequencing Consortium of Alcohol and Nicotine use (GSCAN, N = 1.2 million) (Liu, Jiang, Wedow, Li, Vrieze, et al., 2019) identified 378 single nucleotide-polymorphisms (SNPs) associated with smoking initiation that were conditionally independent at the genome-wide significance level (P-value $< 5 \times 10^{-8}$). Smoking initiation was defined as being an 'ever' vs. 'never' smoker where an 'ever' smoker had to have either smoked 100 cigarettes in their lifetime and/or smoked regularly every day for at least a month. Of the 378 genome-wide significant SNPs, 356 were available in ALSPAC (Liu, Jiang, Wedow, Li, Vrieze, et al., 2019). Considering that smoking is a complex behaviour, of which initiation is only one part, a PRS for lifetime smoking was generated as a sensitivity analysis. The lifetime smoking score also captures

smoking heaviness (as well as smoking duration and cessation) but is derived in the entire population comprising both smokers and non-smokers and therefore is more suitable for use in unstratified samples (Wootton et al., 2019). The GWAS of lifetime smoking based on the UK Biobank data (N = 462,690) identified 126 independent loci (P-value < $5x10^{-8}$), which were all available in ALSPAC. The Coffee and Caffeine Genetics Consortium found eight SNPs to be independently associated with cups of coffee consumed per day at the genome-wide level of significance (N = 91,462) (The Coffee and Caffeine Genetics Consortium et al., 2015), which were all available in ALSPAC. These SNPs have also been found to be associated with caffeine use from other caffeinated beverages (Taylor, Davey Smith, et al., 2018; Treur et al., 2017).

Weighted genetic risk scores were created using independent genome wide significant hits (P-value < $5x10^{-8}$) and their effect estimates as reported in the discovery GWAS for each of the exposures. These PRS were derived using Plink v1.9 and standardised prior to use in analyses. As the PRS were based on discovery GWAS that only report independent variants (Liu, Jiang, Wedow, Li, Vrieze, et al., 2019; The Coffee and Caffeine Genetics Consortium et al., 2015; Wootton et al., 2019), clumping or pruning was not necessary (Choi et al., 2020).

3.3.5 Statistical analysis

All analyses were performed using Stata v15. The following linear and logistic regression analyses were conducted to test associations with the smoking and caffeine PRS: (1) maternal PRS with smoking and caffeine phenotypes in mothers during pregnancy to validate the PRS (Aim 1); (2) maternal and offspring PRS with childhood phenotypes (< 10 years) for investigating intergenerational effects (Aim 2, Figure 3.3). In addition to the two main aims, supplementary analyses were conducted that tested maternal and offspring PRS with their own phenotypes in mothers (during and outside of pregnancy) and offspring (adolescence). Results from the supplementary analyses were used to confirm PRS associations with relevant substance use behaviours outside of pregnancy as a positive control and to gain more information about mental health associations at later times in development. Analyses were adjusted for age, offspring sex, and the first 10 ancestry-informative principal components based on the ALSPAC genome-wide data. The sample was restricted to singletons or one individual from a twin pair and to individuals of European ancestry. The maximum sample size available in childhood was 6,156 (4,974 in adolescence) and 7,269 during pregnancy (7,199 outside of pregnancy). To avoid limiting the sample size further, and to reduce the risk of selection bias, analyses were not restricted to only mother-offspring pairs with complete

genotype data. Differences between smoking, caffeine, and socio-demographic variables were observed between participants with complete vs. partially missing genotype data, which supports the notion that restricting the sample to mother-offspring pairs where both have complete genotype data could have introduced selection bias (Appendix B, Table B1 and Table B2).

3.3.6 Multiple testing

Multiple testing was accounted for by running Monte Carlo permutation testing with 1,000 repetitions. Similar to the nominal P-value that is based on the assumption that the data represents a random set of the general population, a permutation P-value (P_{Perm}) indicates the probability of observing a value that is more or less extreme than the one observed, while taking into account the correlation structure of the analysed data (Conneely & Boehnke, 2007). Permutation testing does not assume that the reference population follows a normal distribution but instead a reference population is generated from random draws (permutations) of the data at hand (Legendre & Legendre, 1998). Permutation testing has been acknowledged as a more adequate approach to correct for multiple testing in PheWAS studies, that aim at generating new hypotheses, than Bonferroni correction (Millard et al., 2015). Bonferroni correction is likely to be overly conservative considering the high degree of correlation between the selected phenotypes (Conneely & Boehnke, 2007; Millard et al., 2015). Thus, permutation testing was applied as the main multiple testing correction and permutation P-values are presented in the text. For phenotypes that showed evidence for association according to a permutation corrected P-value, effect estimates between the intergenerational and childhood analysis were compared. To get an indication for false positive findings, results were also compared with a more stringent Bonferroni correction. Evidence for association was considered strongest for phenotypes that had a permutation P-value of less than 0.05 and also survived Bonferroni correction (all results are available in Appendix C to Appendix I). The analyses for Aim 1, which aimed at validating the caffeine and smoking PRS as proxies for smoking and caffeine consumption during pregnancy, were exempted from the multiple testing corrections, as the validation analyses were hypothesis-testing (Millard et al., 2015).

3.4 Results

3.4.1 Maternal smoking and caffeine consumption

In ALSPAC, 25% of mothers reported smoking during the 1st trimester of pregnancy. Generally, 51% of mothers reported having ever smoked a cigarette in their lifetime, of which 66% reported smoking before pregnancy and 49% reported smoking during the first trimester of pregnancy. Daily caffeine consumption varied widely amongst mothers in ALSPAC and thus the median (and IQR), which is less sensitive to extreme values, is reported instead of the mean and SD. Most mothers in ALSPAC (94%) consumed caffeine outside of pregnancy (8-years post-pregnancy), with a median overall caffeine level of 168 mg/day of caffeine (IQR: 108 to 252) and 38% consuming more than 200 mg/day of caffeine. During pregnancy, most mothers still consumed some caffeine daily (N 2nd trimester = 92%; N 3rd trimester = 90%) but fewer mothers consumed more than 200 mg of caffeine (2nd trimester = 28%; 3rd trimester =, 3rd trimester = 27%). During pregnancy, mothers reported lower caffeine consumption with a median of 138 mg/day (IQR: 81 to 215) during the 2nd trimester and 135 mg/day (IQR: 71 to 216) during the 3rd trimester. Tea was the caffeinated beverage that most mothers consumed outside (82%), as well as during pregnancy (86% during the 2nd, and 88% during the 3rd trimester). Followed by caffeinated coffee (63% outside of pregnancy; and 58% and 53% during the 2nd and 3rd trimester, respectively). Cola was the caffeine source that was least consumed (53% outside pregnancy; and 34% and 41% respectively during the 2nd and 3rd trimester). Stratifying the sample into mothers who reported to consume any caffeine outside of pregnancy (> 0 mg/day caffeine 8-years post-pregnancy) and had data on caffeine consumption during pregnancy (2nd and 3rd trimester, N = 7,191) indicated that across time points in ALSPAC, caffeinated tea was the primary source of caffeine (Figure 3.4). Also, mothers significantly reduced their coffee consumption from the 2nd to 3rd trimester, whereas tea consumption remained relatively stable throughout pregnancy, as well as outside of pregnancy (Figure 3.4). Comparing caffeine consumption during pregnancy between mothers with and without data on caffeine consumption 8-years post-pregnancy showed that mothers with data 8-years post-pregnancy drank on average less caffeine during pregnancy than mothers with missing data 8-years post-pregnancy (2nd and 3rd trimester: mean difference = 19 mg/day caffeine, $t(13{,}289) = 9.17$, P-value < 0.001; $t(12{,}120) = 8.96$, P-value < 0.001), respectively). Likewise for smoking, mothers with data on smoking 8-years post-pregnancy were less likely to have smoked during pregnancy (N = 7,574 of which 19% smoked during pregnancy) than mothers with missing data on smoking 8-years post-pregnancy (N = 5,785 of

which 33% smoked during pregnancy). This indicates that effects of smoking and high caffeine consumption during pregnancy on offspring mental health outcomes (which were assessed in childhood and adolescence several years post-pregnancy) are likely to be underrepresented in this study and effects of high maternal smoking/caffeine consumption might be underestimated (Webb, Bain & Page, 2017).

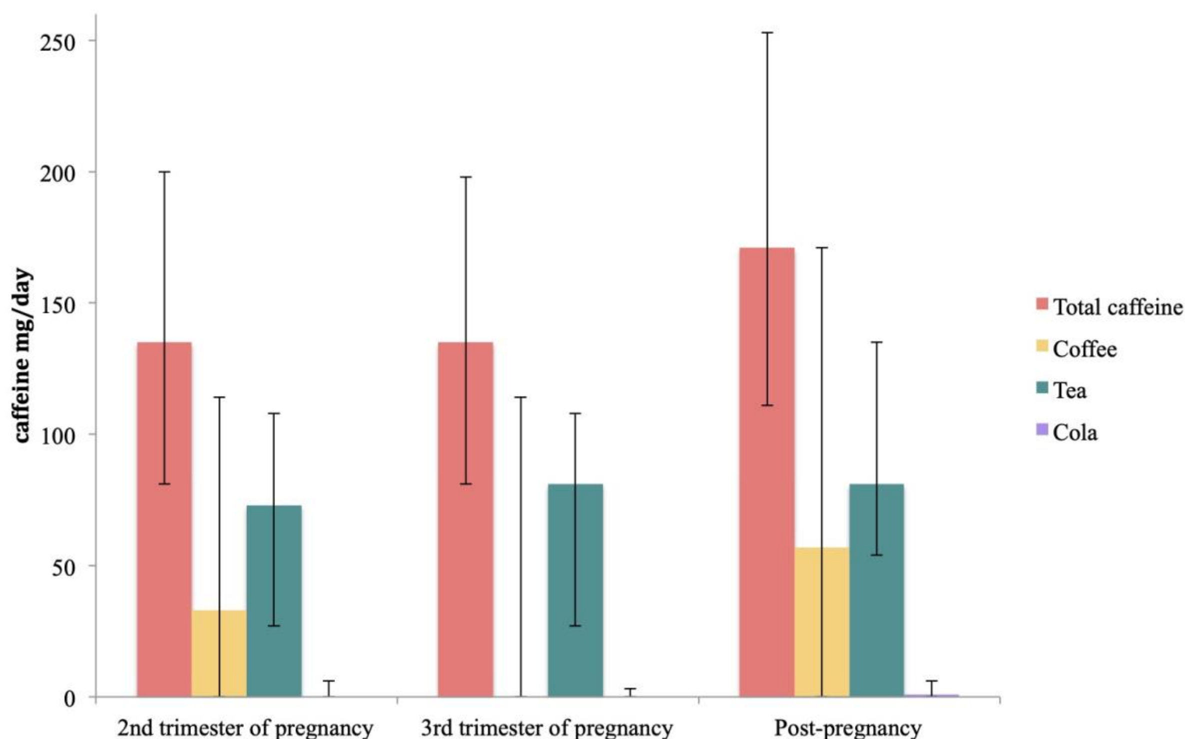


Figure 3.4 Overview of the intake of the different sources of caffeine outside (8-years post pregnancy) and during pregnancy. *Only includes mothers who had complete data on caffeine consumption during pregnancy (2nd and 3rd trimester) and reported consuming some caffeine (> 0 mg/day) 8-years post pregnancy (N = 7,191). Bars represent median mg/day of caffeine per caffeinated drink and error bars represent the IQR.*

Compared to mothers who did not report smoking, mothers who smoked reported consistently more caffeine consumption during (2nd trimester: 64 mg/day more caffeine, 3rd trimester: 75 mg/day more caffeine) and outside of pregnancy (8-years post-pregnancy: 30 mg/day more caffeine).

3.4.2 Aim 1: Validation of PRS during pregnancy

The PRS for smoking initiation and lifetime smoking were positively associated with maternal smoking phenotypes during pregnancy and explained 1-5% of variance in smoking phenotypes during and outside of pregnancy (Table 3.2 and Appendix C, Table C1). The PRS for caffeine consumption was positively associated with total caffeine and caffeinated tea and

coffee consumption during pregnancy, but not with cola consumption (Table 3.3). The caffeine PRS explained 0.2-0.4% of variance in caffeine phenotypes during pregnancy and 0.2-1% of variance outside of pregnancy (Table 3.3). Results of the supplementary positive control analyses, using maternal and offspring smoking and caffeine PRS to predict smoking and caffeine consumption outside of pregnancy and during adolescence, validated that the smoking PRS could predict consumption during these times in life (Table 3.2). The caffeine PRS could predict consumption outside of pregnancy (except for cola consumption) but not during adolescence (Table 3.3).

Table 3.2 Associations between smoking initiation PRS and smoking phenotypes in mothers (during and outside of pregnancy) and offspring in adolescence

	Phenotype	Effect estimate	Effect size*	95% CI	P-value	Sample size	Adj. R^2**
Mothers							
Outside of pregnancy	Mother has ever smoked	OR	1.40	1.33, 1.48	1.24×10^{-8}	7194	0.03
	Number of cigarettes smoked before pregnancy	Beta	0.15	0.08, 0.22	3.81×10^{-5}	3426	0.05
Pregnancy – 18 weeks gestation	Tobacco smoked in 1st three months of pregnancy	OR	1.35	1.23, 1.44	3.0×10^{-7}	7237	0.05
	Mother cut down smoking	OR	1.33	1.25, 1.42	5.89×10^{-7}	7269	0.03
	Mother stopped smoking during pregnancy	OR	0.98	0.88, 1.11	0.771	1863	0.01
Offspring							
Adolescence– 14 years	Smoked at age 14 years	OR	1.18	1.09, 1.28	6.50×10^{-4}	4145	0.03
	Smoked more than 20 cigarettes	OR	1.19	1.03, 1.38	0.024	1058	0.03
	Age 1st smoked a cigarette	Beta	0.001	-0.04, 0.04	0.953	1064	0.01
Adolescence– 18 years	Ever smoked a whole cigarette	OR	1.26	1.15, 1.37	1.09×10^{-4}	2402	0.02
	Number of cigarettes smoked in lifetime	Beta	0.19	0.10, 0.2	4.24×10^{-5}	1144	0.01

Note. * Reflects the average change in the outcome that is associated with a one standard deviation increase in the PRS. For binary outcomes, this will be the odds ratio (e.g., Mother's odds of ever smoking are 1.4 times compared to not smoking), for continuous outcomes it represents the average unit change (e.g., 0.15 cigarettes smoked). ** For the logistic regression models, pseudo R^2 is reported.

Table 3.3 Associations between caffeine PRS and daily caffeine intake in mothers (during and outside of pregnancy), and offspring (age 8 and 13 years)

	Daily caffeine intake phenotype	Effect size* (beta)	95% CI	P-value	Sample size	Adj. R^2**
Mothers						
Outside of pregnancy	Total caffeine	9.89	6.34, 13.44	$4.97x10^{-8}$	4783	0.01
	Coffee	0.03	0.01, 0.06	0.009	4655	0.003
	Tea	0.07	0.03, 0.10	$1.01x10^{-4}$	4632	0.01
	Cola	0.01	-0.01, 0.03	0.332	4670	0.002
Pregnancy – 18 weeks gestation	Total caffeine	5.85	3.09, 8.61	$3.28x10^{-5}$	7220	0.004
	Coffee	0.02	0.01, 0.04	0.011	7198	0.002
	Tea	0.02	0.01, 0.04	0.007	7189	0.002
	Cola	-0.001	-0.02, 0.01	0.890	7185	0.002
Pregnancy – 32 weeks gestation	Total caffeine	6.32	3.74, 8.89	$1.56x10^{-6}$	6767	0.01
	Coffee	0.03	0.01, 0.04	0.01	6596	0.002
	Tea	3.42	1.80, 5.04	$3.65x10^{-5}$	6608	0.004
	Cola	-0.01	-0.03, 0.01	0.278	6500	0.002
Offspring						
Childhood – age 8 years	Total caffeine	0.01	-0.01, 0.03	0.377	4589	0.002
	Coffee	0.01	-0.06,0.08	0.845	254	0.02
	Tea	0.18	-1.52, 1.88	0.836	1475	0.004
	Cola	0.003	-0.02, 0.03	0.829	4551	0.002
Adolescence – age 13 years	Total caffeine	0.01	-0.03, 0.05	0.670	3405	0.004
	Coffee	0.03	-0.02, 0.08	0.271	467	0.05
	Tea	0.89	-0.35, 2.13	0.161	1933	0.004
	Cola	-0.02	-0.05, 0.02	0.424	2411	0.01

Note. * Reflects the average change in the outcome that is associated with a one standard deviation increase in the PRS. For continuous outcomes it represents the average unit change (e.g., a one standard deviation increase in PRS is associated with mothers consuming 9.89 mg/day more caffeine outside of pregnancy). For transformed variables, it represents the average quantile or quartile change (e.g., a one standard deviation change in PRS is associated with a 0.03 quantile mg/day increase in coffee consumption outside of pregnancy). ** For the logistic regression models, pseudo R^2 is reported.

3.4.3 Aim 2: Comparison of intergenerational and childhood smoking initiation PRS analyses

Intergenerational PRS analyses. Of 16 childhood phenotypes, the association with the strongest evidence, according to the smallest permutation corrected P-value, were observed with reduced anxiety symptoms (β_{8years}=-0.03, 95% CI -0.05, -0.01, P_{perm}=0.002) and increased conduct disorder symptoms (β_{7years}=0.02, 95% CI 0.004, 0.04, P_{perm}=0.021). Of the non-mental health phenotypes, the strongest associations, according to the smallest permutation corrected P-value, were found for lower IQ (β_{8years}=-0.59, 95% CI -1.05, -0.134, P_{perm}=0.02), higher overall caffeine consumption (mg/day; β_{8years}=0.05, 95% CI 0.02, 0.07, P_{perm}= < 0.001), and BMI (β_{7years}=0.08, 95% CI 0.02, 0.14, P_{perm}=0.001) as well as the likelihood of being left-handed ($OR_{11years}$=1.11, 95% CI 1.01, 1.23, P_{perm}=0.012), which was included as a negative control phenotype (because a causal intrauterine effect of maternal smoking or caffeine use on handedness would not be expected). Only associations with offspring's anxiety symptoms and caffeine consumption survived Bonferroni correction of P-value < 0.003 (Figure 3.5, Appendix D, Table D1).

Childhood PRS analyses. As observed in the intergenerational analysis, there was some evidence for an association with reduced anxiety problems (β_{8years}=-0.03, 95% CI -0.05, -0.01, P_{perm}=0.002) and increased conduct disorder symptoms (β_{7years}=0.03, 95% CI 0.01, 0.05, P_{perm}=0.001). In contrast to the intergenerational analysis, there was some evidence for an association with ADHD symptoms (β_{7years}=0.03, 95% CI 0.003, 0.06, P_{perm}=0.034). The non-mental health phenotypes that showed the strongest evidence for association in the intergenerational analysis were replicated using the offspring smoking PRS (lower IQ: β_{8years}= -0.74, 95% CI -1.18, -0.29, P_{perm} < 0.001; increased caffeine consumption: β_{8years} = 0.03, 95% CI 0.01, 0.06, P_{perm} = 0.006; BMI: β_{7years}=0.05, 95% CI -0.0003, 0.10, P_{perm}=0.048) with the exception of left-handedness ($OR_{11years}$=1.05, 95% CI 0.95, 1.15, P_{perm}=0.291; Figure 3.5). Only associations with IQ and conduct disorder symptoms survived the Bonferroni correction of P-value < 0.003 (Figure 3.5, Appendix D, Table D1). The results using lifetime smoking PRS were largely consistent. Only associations with offspring's IQ, oppositional defiant disorder (ODD), and total behavioural difficulties survived Bonferroni correction (Appendix E, Table E1).

Comparison of magnitude of effect estimates. For childhood anxiety, the effect estimates were of similar strength in the intergenerational (β_{8years}=-0.03, 95% CI-0.05, -0.01) and childhood analysis (β_{8years}=-0.03, 95% CI -0.05, -0.01), providing no evidence for an intrauterine effect.

The effect estimate for IQ, was slightly larger in the childhood (β_{8years}= -0.74, 95% CI -1.18, -0.29) compared to the intergenerational analysis (β_{8years} =-0.59, 95% CI -1.05, -0.134), indicating that this association is more likely to be pleiotropic than reflecting an intrauterine effect. Even though the effect estimates of the BMI and caffeine phenotypes were slightly larger in the intergenerational (BMI: β_{7years} =0.08, 95% CI 0.02, 0.14; caffeine: β_{8years}=0.05, 95% CI 0.02, 0.07) than the childhood analysis (BMI: β_{7years} =0.05, 95% CI -0.0003, 0.10; caffeine: β_{8years} = 0.03, 95% CI 0.01, 0.06) the overlapping confidence intervals indicate no strong evidence for an intrauterine effect.

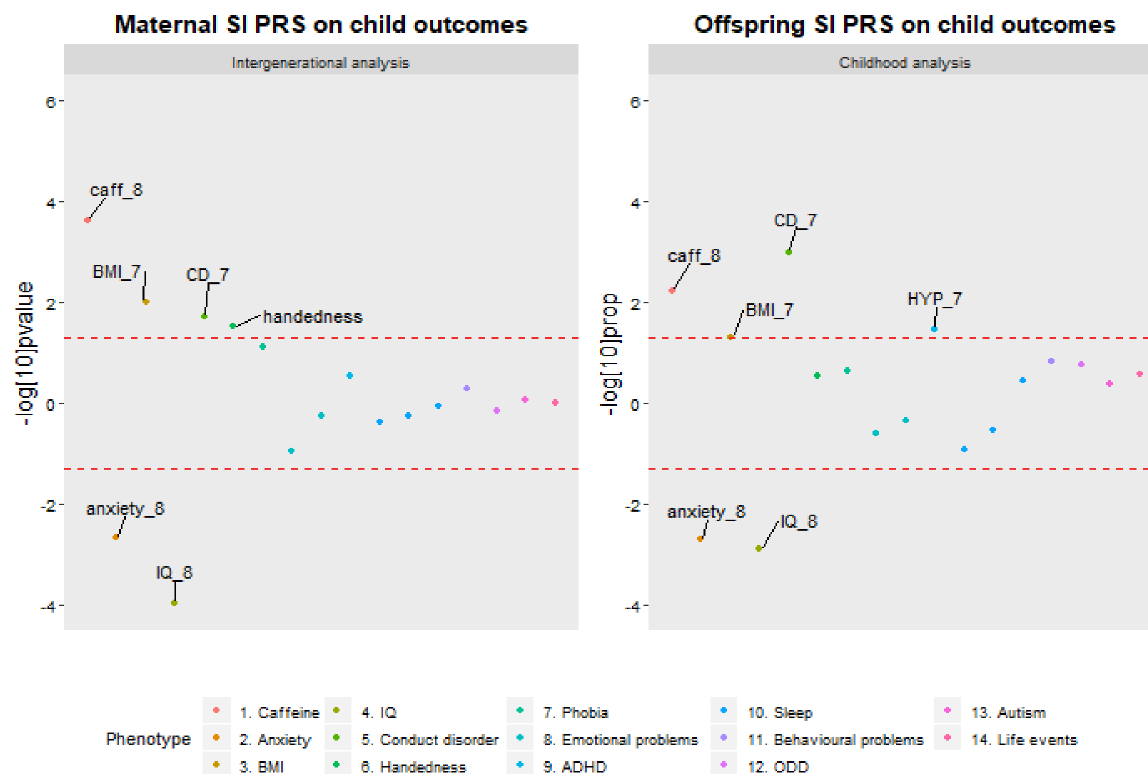


Figure 3.5 Comparison of phenotype associations with the smoking initiation (SI) polygenic risk scores (PRS) in the intergenerational and childhood analysis. *Points outside the lines had a permutation corrected P-value < 0.05. Points above the upper line represent positive associations and points below the lower line represent negative associations. caff_8 = Total caffeine consumption at age 8. BMI_7 = BMI at age 7. CD = Conduct Disorder at age 7. anxiety_8 = Anxiety at age 8. IQ_8 = IQ at age 8. HYP_7 = Hyperactivity at age 8.*

3.4.4 Aim 2: Comparison of intergenerational and childhood caffeine PRS analyses

Intergenerational PRS analyses. Given that offspring's caffeine PRS was not robustly associated with caffeine consumption in childhood (Table 3.3), the results of the childhood analysis could be used as a test for pleiotropy, despite some children already consuming low levels of caffeine at this age. Of the 16 childhood phenotypes, the mental health association with the strongest evidence, according to the smallest permutation corrected P-value, was observed with decreased risk for specific phobias in offspring ($OR_{10years}$ =0.72, 95% CI 0.52, 1.01, P_{perm}=0.028; Figure 3.6 and Appendix F, Table F1). There was no evidence for association with any of the non-mental health phenotypes (Appendix F, Table F1)

Childhood PRS analyses. In contrast to the intergenerational analysis, there was no evidence for association with specific phobias ($OR_{10years}$=1.00, 95% CI 0.72, 1.38, P_{perm}=0.998), but some evidence for association with reduced general anxiety symptoms (β_{8years}= -0.02, 95% CI -0.04, -0.002, P_{perm}=0.026). The association with the strongest evidence amongst the non-mental health phenotypes, according to the smallest permutation corrected P-value, was observed with fewer hours of sleep in term-time (β_{7years}= -0.03, 95% CI -0.05, -0.004, P_{perm}=0.018), (Figure 3.6, Appendix F, Table F1). None of the associations of the intergenerational and childhood analyses for caffeine survived Bonferroni correction.

Comparison of magnitude of effect estimates. The effect estimate of the association between the caffeine PRS and risk for specific phobias was larger in the intergenerational ($OR_{10years}$ =0.72, 95% CI 0.52, 1.01) compared to the childhood analysis ($OR_{10years}$=1.00, 95% CI 0.72, 1.38). As the confidence intervals were partially, but not completely, overlapping this might indicate an intrauterine effect of caffeine on reduced risk for offspring specific phobias, or chance. For general anxiety symptoms there was no evidence for an intrauterine effect, as the childhood analysis showed a similar size of effect (β_{8years}= -0.02, 95% CI -0.04, -0.002) as the intergenerational analysis (β_{8years} = -0.02, CI -0.02 to 0.02). The phenotype of fewer hours of sleep in term-time showed a similar strength of association in the childhood (β_{7years}= -0.03, 95% CI -0.05, -0.004) and intergenerational analysis (β_{7years}= -0.03, 95% CI -0.03, 0.02) also providing no evidence for an intrauterine effect.

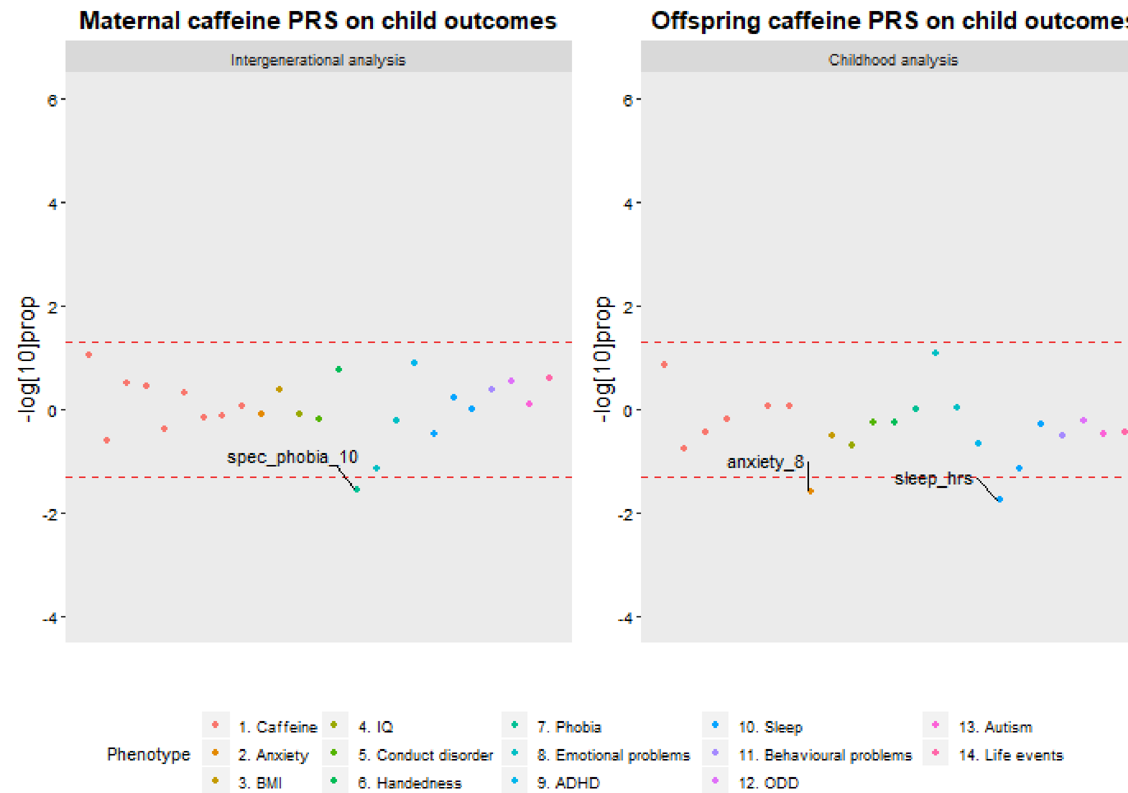


Figure 3.6 Comparison of phenotype associations with the caffeine polygenic risk scores (PRS) in the intergenerational and childhood analysis. *Points outside the lines had a permutation corrected P-value < 0.05. Points above the upper line represent positive associations and points below the lower line represent negative associations. spec_phobia_10 = specific phobias at age 10. anxiety_8 = Anxiety at age 8. sleep_hrs = Sleep duration in hours at age 7.*

3.4.5 Supplementary analyses: Associations between maternal and offspring PRS and own mental health during and outside of pregnancy (mothers) and adolescence (offspring).

In addition to the intergenerational and negative control analyses, associations between the maternal and offspring PRS and *own* mental health during and outside of pregnancy (mothers) and adolescence (offspring) were explored for pleiotropic or suggestive causal relationships. Further, by comparing the magnitude of effects across childhood, adolescence, and adulthood, the persistence of pleiotropic effects across development and the effect on mental health through maternal and offspring's own behaviour could be explored.

3.4.5.1 Maternal smoking initiation PRS and outcomes during and outside of pregnancy.

Amongst the mental health phenotypes, there was consistent evidence for the maternal smoking PRS being associated with increased depressive symptoms during and outside of pregnancy. After applying the Bonferroni correction (threshold: P-value < 0.002), the strongest evidence was found for associations with lower education, higher caffeine consumption, and binge drinking during and outside of pregnancy, as well as with higher BMI, lower image perception, and personality traits such as more anger and monotony avoidance outside of pregnancy (Appendix G, Table G1).

3.4.5.2 Offspring smoking initiation PRS and own outcomes during adolescence

Consistent with the results using the childhood mental health phenotypes, there was some evidence for positive associations with externalising problems in adolescence (conduct disorder, oppositional defiant disorder, ADHD). However, there was very little evidence for associations with anxiety symptoms during adolescence. After applying the Bonferroni correction (threshold: P-value < 0.001), the strongest evidence was found for associations with increased conduct disorder symptoms, higher BMI, lower IQ, more extraverted personality traits, and alcohol consumption (Appendix G, Table G1).

3.4.5.3 Offspring lifetime smoking PRS and own outcomes during and outside of pregnancy (mothers) and during adolescence (offspring)

The results using the lifetime smoking PRS were largely consistent and can be found in Appendix H (Table H1 and Appendix C, Table C1). The strongest evidence was found for associations with education, social class, caffeine consumption, BMI, and anger phenotypes in mothers during and outside of pregnancy and conduct disorder, extraversion, and IQ phenotypes in adolescence (Appendix H, Table H1).

3.4.5.4 Maternal caffeine PRS and own outcomes during and outside of pregnancy

There was some evidence for association with decreased likelihood of having had schizophrenia diagnosis outside of pregnancy but no other mental health outcomes during or outside of pregnancy (Appendix I, Table I1). Amongst the non-mental health phenotypes, some evidence was observed for associations with less substance use during pregnancy (higher likelihood of reducing caffeine consumption, decreased likelihood to binge drink), as well as evidence for an association with lower socio-economic position. None of these associations survived the Bonferroni correction neither outside (threshold: P-value < 0.002) nor during pregnancy (threshold: P-value < 0.001; Appendix I, Table I1).

3.4.5.5 Offspring caffeine PRS and own mental health outcomes during adolescence

The caffeine PRS was not associated with caffeine consumption during adolescence (Table 3.3). There was only some evidence for the offspring caffeine PRS being associated with higher GCSE exam grades during adolescence but none of the mental health or substance use phenotypes. None of these associations survived Bonferroni correction (threshold: P-value < 0.001; Appendix I; Table 0).

3.5 Discussion

3.5.1 Summary and interpretation of findings

The aim of this study was to identify possible causal associations of maternal smoking and caffeine consumption on offspring mental health, which could be followed-up in analyses with a particular focus on the prenatal period. The results showed that the smoking and caffeine PRS are valid predictors of smoking and caffeine consumption from tea and coffee during pregnancy. The smoking initiation PRS was associated with various psychological traits and other substance use behaviours across different time points in life. In particular, associations were observed between the smoking initiation PRS and sensation-seeking traits across development, such as less anxiety and increased externalising problems in childhood, an extroverted personality type, more externalising problems and alcohol consumption in adolescence, as well as higher expression of anger, more monotony avoidance outside of pregnancy and alcohol consumption during and outside of pregnancy. The caffeine PRS showed weak evidence for associations with offspring mental health outcomes. For the phenotypes that showed evidence for an association with the smoking and caffeine PRS, no strong difference in effect estimates could be observed between the childhood and intergenerational analysis, providing no evidence for an intrauterine effect. Critically, the results indicate evidence that the associations found between the maternal smoking and caffeine PRS and offspring mental health outcomes are likely due to pleiotropic effects, rather than acting through the maternal intrauterine environment.

3.5.2 Smoking PRS and mental health

The findings of pleiotropic associations between the smoking PRS and sensation-seeking personality traits are supported by the literature. Observational evidence suggests that adolescents who start smoking tend to have more externalising behavioural problems (Crone & Reijneveld, 2007; Elkins et al., 2007; Pedersen et al., 2018), are more novelty seeking, and have lower cognitive abilities, socio-economic position, and academic outcomes (Audrain-McGovern et al., 2004; Daly & Egan, 2017; Pedersen et al., 2018). Results of genetic analyses suggest that a shared genetic liability for smoking, externalising problems, and impulsivity (Harrison et al., 2019; Hicks et al., 2020; Stephens et al., 2012; Young

et al., 2000), as well as socioeconomic variables (education and SEP; Gage et al., 2018; Khouja et al., 2021), may explain observational associations (Iacono et al., 2008). There was some evidence for the smoking PRS being associated with maternal depression during and outside of pregnancy, which could (partly) explain the association observed between the smoking PRS and offspring externalising problems (Eilertsen et al., 2020). A study adopting a similar design to the present one, examining associations between maternal and offspring PRS for increased alcohol consumption and maternal and offspring mental health, also found an association between maternal alcohol use and maternal depression during pregnancy but no evidence for an association with maternal alcohol PRS and externalising problems in offspring (Easey et al., 2021). Even though this requires further testing, it could provide some initial evidence that the association between the smoking PRS and offspring externalising problems is more likely to be pleiotropic than confounded by maternal depression. Especially since associations with impulsivity and sensation-seeking behaviours, as well as ADHD, were still evident in studies that only included genome-wide significant SNPs (Harrison et al., 2019; Hicks et al., 2020; Khouja et al., 2020; Liu et al., 2019). Contextualising the results of this study with the abovementioned literature suggests that associations observed with the smoking initiation PRS are reflecting a common genetic liability with sensation-seeking phenotypes that is shared between mother and offspring.

3.5.3 Caffeine PRS and Mental Health

The caffeine PRS did neither predict maternal nor offspring's caffeinated cola consumption, which could be explained by low consumption of cola in our sample and/or the low caffeine content in cola compared to coffee and tea. Further, it is possible that cola has a different underlying genetic or confounding structure than coffee and tea. During adolescence, the coffee PRS was not associated with any of the three included sources of caffeine. This in contrary to previous research that compared the development of caffeine consumption between monozygotic and dizygotic twins and found that genetic influences on caffeine consumption increase between the ages 9 to 13 years of age, and then remain stable at explaining 30% to 45% of variance in caffeine consumption (Kendler et al.,

2008). However, the sample of that study was quite different from the ALSPAC sample (only included males born between 1940-1974). Thus, the possibility remains that no genetic effect was observed for adolescents in this study due to low average caffeine consumption at the assessment point of 13 years (overall caffeine: $Median_{13years}$ = 20 mg/day, $IQR_{13years}$: 3 to 43). This is in line with findings from The Coffee and Caffeine Genetics Consortium, in which the PRS for coffee consumption was not replicated in a Pakistani population, where coffee consumption is rare (The Coffee and Caffeine Genetics Consortium et al., 2015). Also, it has been argued that the genetic contribution to caffeine consumption is stronger for heavier than modest caffeine consumers (Yang et al., 2010). However, as the discovery GWAS did not include adolescent populations, it is also possible that there are different genetic variants contributing to caffeine consumption during adolescence compared to adulthood. In sum, as the caffeine PRS could not predict caffeine consumption during adolescence, the associations between the caffeine PRS and mental health outcomes assessed in adolescence may not represent effects of offspring's caffeine consumption on adolescent mental health.

Results did not show strong evidence for intergenerational effects between the maternal caffeine PRS and offspring mental health outcomes in childhood. There was some evidence for pleiotropic associations of the caffeine PRS with lower risk for anxiety symptoms. Again, this finding ties in with other research stating pleiotropic effects of more risk-prone personality traits and increased substance use (Khouja et al., 2021; Malmberg et al., 2010). Further, previous research suggests that people with a predisposition towards anxiety consume less caffeine because it increases their anxiety symptoms and that caffeine does not have the same effect in less anxious people (Lara, 2010; Lee et al., 1985). The associations that were observed between caffeine PRS and decreased likelihood of binge drinking, reduced caffeine consumption, and lower socioeconomic position during pregnancy, as well as higher General Certificate of Secondary Education (GCSE) exam grades during adolescence stand in contrast to a study in the UK Biobank, where the caffeine PRS was positively associated with alcohol consumption outside of pregnancy and not associated with social class (Taylor, Davey Smith, et al., 2018). Therefore, these findings should be interpreted with caution, as they

might be unique to the ALSPAC sample in terms of participants' sociodemographic characteristics or false positives. Although these results could be due to yet unexplained forms of bias, it is also possible that the caffeine PRS is capturing underlying personality/socio-behavioural profiles with far reaching consequences for health and wellbeing, which deserves further investigation.

3.5.4 Smoking and caffeine consumption

The descriptive statistics of maternal caffeine consumption confirm what has been found in previous studies with mothers generally reducing their overall caffeine intake during pregnancy, specifically consuming less caffeinated coffee and slightly more caffeinated tea (Figure 3.3) (Chen et al., 2014; Lawson et al., 2004) and mothers who smoke, consuming more caffeine during pregnancy than mothers who do not smoke during pregnancy (Chen et al., 2014; Loomans et al., 2012; Robinson et al., 2010).

Both the smoking initiation and lifetime smoking PRS showed evidence for associations with increased maternal caffeine consumption outside and during pregnancy (Appendix G, Table G1 and Appendix H, Table H1). However, only the smoking initiation PRS also showed evidence for associations with increased caffeine consumption in the childhood and intergenerational analyses (Figure 3.5; Appendix D, Table D1). Previous research suggested a shared genetic liability between smoking and caffeine consumption (Kendler et al., 2008; Treur et al., 2017), however, children consume quite low levels of caffeine before the age of 10 and even the caffeine PRS was not able to predict caffeine consumption during childhood. Furthermore, the smoking initiation PRS was not associated with caffeine consumption during adolescence (Appendix G, Table G1). The possibility remains that associations detected in the intergenerational and childhood analysis are (partly) reflecting effects of genetic nurturing. Genetic nurturing occurs when genetic variants that are used as a proxy for an environmental exposure also influence the raring environment of the child (Pingault et al., 2018). Therefore, it is possible that the smoking initiation instrument captures variance of a latent construct that influences childhood caffeine consumption. For instance, parenting practices have been found to differ between mothers who smoked and did not smoke during pregnancy (Tandon et al., 2009) and considering that the smoking PRS is likely to capture risk-prone

personality traits, parents with a higher smoking initiation PRS might tend to be less cautious about whether their children consume caffeinated beverages. Also, due the comorbidity of smoking and caffeine consumption, it is likely that children born to parents with a higher smoking initiation PRS might in general be more exposed to caffeinated beverages in their home environment. Adding to this, offspring with a higher smoking initiation PRS may be more sensation seeking and thus more likely to try caffeinated drinks as a child. On the other hand, the association could also represent a false positive finding and requires further testing in independent samples, integrating MR with family-based designs (Pingault et al., 2018).

3.5.5 Strengths and Limitations

A major strength of this study was the exploration of exposure-outcome associations at time points in life other than adulthood. Further, the validation of genetic variants discovered in non-pregnant female and male populations, as proxies during pregnancy, is vital for future investigation of intrauterine effects of the exposures (Lawlor et al., 2017). Lastly, the intergenerational comparison of associations of the maternal smoking and caffeine PRS with childhood mental health outcomes, that are likely to be free of confounding through offspring's own substance consumption enabled to disentangle potential pleiotropic and environmental effects on mental health.

The following limitations should also be considered. First, the limited sample size (in the context of genetic association studies) likely resulted in low statistical power to detect small effects. This may also have compromised the precision of effect estimates and thus no strong inferences can be made based on the comparison of effect estimates between the childhood and intergenerational analyses. Second, mental health outcomes were restricted to phenotypes as assessed in ALSPAC, and the comparison of related phenotypes was not similar across development (e.g., ADHD/conduct disorder in childhood with extraversion & anger personality traits in mothers outside of pregnancy). Third, many mental health phenotypes in childhood were based on maternal report, which may not accurately reflect offspring's mental health problems (Gartstein et al., 2009; Najman et al., 2001) but rather mothers own mental health status (Hennigan et al.,

2006; Ringoot et al., 2015). Fourth, PRS for smoking initiation were based on the latest GWAS that included ALSPAC mothers (Liu et al., 2019). Due to the sample overlap, the true strength of explored associations might be smaller than reported. However, given the small contribution of data from ALSPAC (~1%) to a total sample size of 1.2 million, the risk of bias is likely negligible. Fifth, to make the smoking PRS specific to our exposure of interest, the PRS was based on genome-wide significant SNPs only, yet the smoking PRS still showed associations with some alcohol phenotypes. Correlations were checked between the alcohol, smoking and caffeine PRS, which were low (Appendix J, Table J1). However, because of the phenotypic associations with alcohol consumption, it cannot rule out that associations observed with the maternal smoking PRS are cofounded by maternal alcohol consumption. Still, this is unlikely to affect the results of this study because there was no evidence for potential causal effects, and previous research by Easey and colleagues (Easey et al., 2021) observed no associations in intergenerational analyses between maternal alcohol PRS and offspring mental health outcomes (Easey et al., 2021). Sixth, as the dataset included phenotypes from later time points and only participants whose genotype data was available, it is possible that the findings are subject to selection bias (see Chapter 1) (Munafò et al., 2018; Taylor, Jones, et al., 2018). This has also been indicated by the comparison of smoking status and caffeine consumption during pregnancy between mothers with and without missing data 8-years post-pregnancy (see section 3.4.1). Last, the comparison of the intergenerational and childhood PRS analyses was based on transmitted alleles and therefore an indirect effect of maternal non-transmitted alleles on offspring sensation-seeking traits through genetic nurturing cannot be ruled out (Kong et al., 2018).

3.5.6 Future research

Future studies investigating the effects on mental health using the smoking initiation PRS might consider accounting for sensation seeking personality traits. Further, future research should aim to differentiate effects of smoke exposure through the intrauterine and postnatal environment, explore non-linear effects of the smoking and caffeine PRS, and investigate a potential interaction of smoking and caffeine consumption during pregnancy on offspring mental health (Grosso & Bracken, 2005). More analyses exploiting paternal data would be helpful to

understand the effect of smoking and caffeine consumption on offspring mental health outcomes. For instance, studies with paternal genotype data could help to differentiate whether observed effects are due to intrauterine or postnatal exposure, through conducting negative control comparisons of prenatal associations of maternal and paternal substance use. Last, studies may want to follow-up the associations that were observed between the maternal smoking PRS and offspring metal health outcomes by integrating MR analyses in family-based designs to account for potential effects of genetic nurturing (Pingault et al., 2018).

3.5.7 Conclusion

In conclusion, this study validated the application of the smoking initiation, lifetime smoking and caffeine PRS for research investigating intrauterine exposures to smoking and caffeinated coffee and tea. Further, results of this study show stronger evidence for pleiotropic than causal effects of maternal smoking and caffeine consumption on offspring mental health (particularly externalising problems). Research using the smoking initiation PRS for investigating effects of smoking on externalising and internalising problems in children should be wary that, due to potential pleiotropy with sensation seeking personality traits, effects on externalising problems might be overestimated, whereas effects on internalising problems might be underestimated. Given the current study's limitations, particularly its limited statistical power, these findings should be replicated in independent and larger samples using more refined methods for pleiotropy detection and corrections.

Chapter 4 – Maternal caffeine consumption during pregnancy and offspring cord blood DNA methylation: a meta-analysis of six epigenome-wide association studies

4.1 Chapter overview

As mentioned in Chapter 1, estimating causal effects of caffeine use during pregnancy on offspring health purely based on conventional observational studies is highly problematic due to problems of confounding. Even though the results of the phenome-wide association study (PheWAS; Chapter 3) did not indicate an effect of maternal caffeine consumption on offspring mental health outcomes, small to moderate effects of maternal caffeine consumption on offspring mental health might have been missed due to statistical power limitations. Furthermore, the analysis was restricted to caffeine consumption in a UK sample and results did not uniquely inspect effects of caffeine consumption *during* pregnancy. Therefore, this chapter specifically focusses on the effect of maternal caffeine consumption during pregnancy on offspring. As elaborated in Chapter 1, offspring DNA methylation is a potential molecular mediator of the relationship of caffeine on offspring mental health outcomes and could provide evidence for a causal intrauterine effect of caffeine. Therefore, I planned and conducted an EWAS meta-analysis study to explore the extent to which maternal caffeine consumption (from tea, coffee, and cola) during pregnancy is associated with offspring cord blood DNA methylation across the genome (step 2 of the meet-in-the middle approach: association between exposure and potential biomarker of the exposure). As caffeine consumption greatly differs culturally (see Chapter 1), the EWAS meta-analysis included cohorts from different European countries to reduce culture-specific confounding. I also conducted additional analyses (using paternal caffeine consumption as a negative control and a polygenic risk score (PRS) to proxy for maternal coffee consumption) in an attempt to infer whether identified associations were causal (e.g., specific to caffeine, as opposed to due to confounding factors/behaviours).

4.2 Introduction

Animal studies have shown that prenatal consumption can lead to changes in DNA methylation and offspring risk of disease (Buscariollo et al., 2014; Fang et al., 2014; Wu et al., 2015; Xu et al., 2012). In humans, a recent large epigenome-wide association study (EWAS) meta-analysis, including 15 international cohorts, found associations between own coffee and tea consumption and (peripheral blood) DNA methylation differences at 11 CpG sites (Karabegović et al., 2020), but I have found no published studies of maternal caffeine consumption during pregnancy in relation to offspring DNA methylation. Considering the strong overlap between smoking and caffeine consumption, it cannot be ruled out that the strong associations found between maternal smoking during pregnancy and changes in offspring DNA methylation (Joubert et al., 2016), are partly attributable to effects of maternal caffeine consumption.

Caffeine consumption during pregnancy tends to be positively associated with other lifestyle behaviours, such as smoking and alcohol consumption, and negatively associated with maternal age and socio-economic position (Chen et al., 2014). Coffee consumption, in particular, is strongly associated with smoking (Sengpiel et al., 2013; Swanson et al., 1994), which can heavily confound associations with health outcomes. The association of smoking with consumption of other caffeinated drinks (e.g., tea and cola) is less clear, with some studies reporting positive associations (Sengpiel et al., 2013; Treur et al., 2016) and others reporting negative or null associations (Swanson et al., 1994; Taylor, Davey Smith, et al., 2018). Potential reasons for the conflicting results might be the lower caffeine content in these beverages and/or different patterns of confounding (Bjørngaard, Nordestgaard, Taylor, Treur, Gabrielsen, Munafò, Nordestgaard, Åsvold, Romundstad, & Davey Smith, 2017; Treur et al., 2016). For instance, an international analysis investigating associations between smoking and caffeine consumption from coffee, tea, and cola, found evidence for cultural confounding of tea (Treur et al., 2016). Whereas a positive association between smoking and tea consumption was found in the British birth cohort, the association with smoking was reversed in the Dutch cohort. Adding to this, tea consumption was associated with higher socio-economic position and education and lower alcohol consumption and BMI in the Dutch but not in the British birth

cohort, suggesting that confounding effects are context- and culture-specific. To better understand the effects of intrauterine caffeine exposure, it is important to study and compare associations with caffeine from different sources and different contexts for the same source.

4.3 Methods

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

4.3.1.1 Participating cohorts

The EWAS meta-analysis included six independent prospective pregnancy and birth cohorts from the PACE consortium (Felix et al., 2018), that had data on cord blood DNA methylation and maternal caffeine consumption during pregnancy available. The total sample (N = 3,731) included two UK based cohorts (ALSPAC and BiB), one Dutch cohort (Generation R), one Norwegian (MoBa1), one Spanish (INMA), and one French (EDEN) cohort. Recruitment periods varied by cohort and took place between the beginning of the 1990s (ALSPAC) and 2009 (MoBa1 and INMA). More details about the individual cohorts can be found in the cohort section (2.6) of Chapter 2.

4.3.1.2 Measurement of maternal caffeine intake during pregnancy

Assessment of maternal caffeine consumption varied by cohort and is described in more detail in Appendix K, K1 to K6. Generally, mothers self-reported the number of cups they consumed of caffeinated coffee, tea, and cola in questionnaires between weeks 12 to week 22 of pregnancy. All cohorts used Food Frequency Questionnaires (FFQ) (Thompson & Subar, 2017) 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 (Farrow et al., 1998). A continuous total caffeine score was calculated by summing the caffeine content from each caffeinated drink in mg/day and allowing for partially missing data by only excluding participants if they had missing data on all three caffeine variables (e.g., if a mother reported to consume two cups of tea per day but had missing data for cups of coffee and cola, she would be given a total caffeine score of 2 x 27 = 54 mg/day of caffeine). Further,

to control for data entry errors, participants were excluded if they reported consuming more than 5 standard deviations from the mean (equivalent to roughly 5 x 82mg/day = 410 mg/day caffeine, which is approx. 7 cups of coffee or 15 cups of tea). In addition to the continuous score, I investigated whether any caffeine exposure (regardless of the amount of caffeine) during pregnancy might influence offspring DNA methylation, and thus dichotomised total caffeine into 0 = none, and > 0 = any.

4.3.1.3 Measurement of DNA methylation

Cord blood DNA methylation was assessed through normalised beta values ranging from 0 to 1, representing 0 to 100% methylation. Cohorts assessed methylation data individually, using their own laboratory methods, quality control, and normalisation. DNA methylation data was sampled using the lllumina Infinium® HumanMethylation450 (486,425 probes), except for BiB, which used the EPIC BeadChip array (866,553 probes). Probes on single nucleotide polymorphisms (SNPs), cross-hybridizing probes (Chen et al., 2013), 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.

4.3.1.4 Other covariates

To adjust for biological variation in DNA methylation, models were adjusted for estimated cell proportions using the Houseman method with a cord blood reference panel (see Chapter 2 for more detail) (Gervin et al., 2016; Houseman et al., 2012), as well as a binary measure of offspring's sex that was used as a stratification factor in a sex-stratified sensitivity analysis. As highlighted in Chapter 2, there is indication that CpG sites associated with gestational age may be specifically prone to changes in DNA methylation (Xu et al., 2017) and thus, gestational age was included as a sensitivity analysis in a separate model. To adjust for possible technical variation (explained in more detail in Chapter 2), all cohorts generated 20 surrogate variables and included them in models (as standard practice in the field: Sharp et al., 2021).

In addition to the necessary biological and technical confounders that need to be adjusted for in EWAS analyses (see Chapter 2), I included the following variables as covariates that are commonly adjusted for in EWAS analyses of maternal

exposures (Joubert et al., 2016; Sharp et al., 2021), because they might influence offspring DNA methylation: (1) An ordinal measure representing maternal education as a proxy for socio-economic position (Alfano et al., 2019), (2) maternal age in years (Markunas et al., 2016), maternal BMI (Sharp et al., 2017), a binary measure of maternal smoking during pregnancy (e.g., in ALSPAC: 0 = no smoking/ giving up smoking during the first trimester, 1 = smoking after the 1st trimester) (Joubert et al., 2016). Last, a binary assessment of parity (1= one or more previous children, 0 = no previous children) was included, because compared to women who have given birth before, women who are pregnant for the first time were found to experience more stress during pregnancy (Gillespie et al., 2018; Lynn et al., 2011). Further, there is evidence that the intrauterine environment changes after the first pregnancy (Ballering et al., 2018) and that the likelihood for adverse birth outcomes is higher for nulliparous than multiparous women (Shah, 2010). See Appendix K, K1 to K6 for the classifications of covariates in each cohort.

4.3.2 Statistical analyses

4.3.2.1 Cohort-specific statistical analyses

4.3.2.1.1 Probe-level analysis

I generated an analysis plan and R script that were made available on GitHub (https://github.com/ammegandchips/Prenatal_Caffeine). I asked cohorts 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, I asked them to run the EWAS analysis separately for each group. The EWAS R script included: a function to exclude measurement error of DNA methylation probes, by removing probes that exceeded 3 times the IQR (Tukey, 1977), a function to generate surrogate variables using the R package SVA (Leek et al., 2019), and a function to run an EWAS of each model using the R package Limma (Ritchie et al., 2015). For each model, the EWAS function ran a linear regression model at each CpG site using maternal caffeine phenotypes as the exposure and offspring cord blood DNA methylation as the outcome. For each of the different caffeine phenotypes (any vs. no caffeine; total caffeine, caffeine from coffee, tea, and cola), I adjusted for covariates (maternal education, maternal age,

maternal BMI, maternal smoking during pregnancy, parity, estimated cell counts). In a sensitivity analysis, I stratified the any/none and total caffeine models by offspring's sex to account for variation due to differential exposure to androgens during pregnancy (Suderman et al., 2017). In a second sensitivity analysis, I additionally adjusted for gestational age at birth. The rationale for this is that gestational age is robustly associated with DNA methylation (Merid et al., 2020; York et al., 2019) and there is some evidence that it can be associated with prenatal caffeine exposure (Bakker et al., 2010; Hoyt et al., 2014). However, it is not likely to be a true confounder, but rather could be a mediating factor on the causal pathway, so I wanted to avoid adjusting for it (which might introduce collider bias) in the main models.

In addition to the EWAS analysis, I hypothesised that intrauterine caffeine exposure might influence cell composition in offspring cord blood (see Chapter 2, section 2.3.2.2), so I conducted linear regression analyses using the main caffeine phenotypes (any vs. no caffeine and total maternal caffeine consumption) as the exposures and estimated offspring cord blood cell proportions as the outcomes.

4.3.2.1.2 Quality control checks for cohort results

Prior to meta-analysing summary results from each cohort, I conducted quality checks to ensure that the EWAS were properly conducted and there were no problems with the data, in line with standard practice in the field (Sharp et al., 2020; Van der Most et al., 2017). First, I generated correlation matrices of the beta coefficients for each of the models. Second, I plotted the distributions of the P-values in QQ-plots and calculated Lambda values. Further, I generated precision plots by plotting 1/median standard error against the square root of the sample size of each cohort.

4.3.2.1.3 Differentially methylated regions (DMR)

I complemented the probe level approach 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 makes more sense biologically; neighbouring CpG sites are assumed to exert similar biological functions. I used the DMRff method (Suderman et al., 2018) to identify differentially methylated regions (DMRs). I supplied cohorts with an R script to

conduct the DMR analysis using their own data. The code used the *dmrff.pre* function, which generates an R-object that includes each cohort's EWAS summary statistics and a sparse correlation matrix from the cohort's methylation data (Suderman et al., 2018). Probes were annotated to the human reference genome version 19, build 37h using the annotation data available from the R-package *meffil* (Suderman et al., 2019). Cohorts sent me the objects that they generated using the *dmrff.pre* function for the DMR meta-analysis.

4.3.2.2 Meta-analysis

4.3.2.2.1 Probe level meta-analysis

Because I was presuming that each of the cohorts EWAS was estimating the same effect, I meta-analysed results with fixed effect estimates weighted by the inverse of the variance using the software METAL (Willer et al., 2010). To correct for multiple testing, I calculated an FDR-adjusted P-value to determine statistical significance below the nominal P-value threshold of 0.05 (Benjamini & Hochberg, 1995).

4.3.2.2.2 Quality control checks for meta-analysed results

I scrutinized the meta-analysis results in a similar manner as the individual cohort results. I generated correlation matrices of effect estimates of the meta-analysed models and investigated the distributions of P-values through QQ-plots and generation of Lambda values. Additionally, I performed a leave-one-cohort-out analysis using the R package *metafor* (Viechtbauer, 2010), which allowed me to inspect how the meta-analysis effect estimates are affected by removal of one cohort (with each cohort being removed sequentially). I deemed results 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.

4.3.2.2.3 DMR meta-analysis

I applied the *dmrff.meta* function (Suderman et al., 2018) to the R-objects that the cohorts sent me to perform the DMR meta-analysis. The dmrff.meta function first determines prospective candidate DMRs by screening regions in the in the meta-analysed EWAS results for consecutive CpG sites that show the same direction of

effect, P-values smaller than 0.05, and a gap of at least 500 base-pairs between consecutive CpG sites (Suderman et al., 2018). Second, for these candidate DMRs, summary statistics are calculated within each cohort by meta-analysing the cohort-level EWAS summary statistics of the CpG sites in the region using a variation of the inverse weighted fixed effects meta-analysis approach that accounts for dependence between CpG sites. Last, the DMR summary statistics of each cohort are meta-analysed to obtain the final cross-cohort meta-analysed DMR statistics and P-values. The cross-cohort meta-analysis uses an inverse-variance weighted fixed effects approach within the dmrff.meta function and results are adjusted for the number of conducted tests using the Bonferroni correction. In this study, I defined a DMR as a region with at least two CpG sites with the same direction of effect and a Bonferroni adjusted P-value ($P_{Bonferroni}$) < 0.05.

4.3.2.3 Causal inference and sensitivity analyses

I conducted the following analyses in an attempt to infer whether associations at the identified CpG sites (from either the probe-level or region-based meta-analyses) were likely to reflect causal intrauterine effects of maternal caffeine (i.e., be specific to caffeine) on offspring cord blood DNA methylation, or rather be spurious associations caused by confounders or other biases.

4.3.2.3.1 Beverage-specific effects across meta-analysed caffeine models

I investigated beverage-specific effects of the meta-analysed probe-level and DMR results by comparing the congruence between associations found using different sources of caffeine (that could have different confounding structures). If caffeine would be having a causal effect on offspring DNA methylation, I would not expect to find evidence for beverage-specific effects. Instead, I would expect models, using different sources of caffeine, to show high congruence in their results (in terms of CpG site hits and/or genes annotated to CpG sites found in the different models), despite potentially having different confounding structures. I expected beverage-specific effects to be indicated by low congruence of results between the different caffeine models, which would provide evidence that instead of caffeine driving the associations (which is common to all beverages), other factors than caffeine are driving the associations.

4.3.2.3.2 Negative control analysis: Paternal caffeine intake (ALSPAC only)

To explore whether any of the maternal caffeine associated CpG sites are indicative of an intrauterine effect of caffeine on DNA methylation, I ran a negative control analysis using paternal caffeine phenotypes as the exposure and the CpG sites that showed the strongest evidence for association with maternal caffeine consumption as the outcome. In theory, even if maternal and paternal caffeine consumption are highly correlated, a true intrauterine effect of caffeine should be reflected by a larger effect estimate in the maternal caffeine analysis than the effect estimates in the paternal caffeine analysis (Taylor et al., 2017). However, if the effects were equally strong in the maternal and paternal caffeine analysis, this would provide some evidence for the maternal associations being confounded rather than representing an effect through the intrauterine environment (based on the assumption that maternal and paternal caffeine consumption underlie a similar confounding structure, but only the former has a plausible biological effect on the outcome through intrauterine exposures).

Paternal caffeine consumption phenotypes (from coffee, tea, cola, total, any/none) were derived as described for maternal caffeine and used as exposure variables in linear regression models with DNA methylation as the outcome. All models were adjusted for the covariates listed above and estimated cell counts. Four models were compared: (1) paternal caffeine with no further adjustment, (2) paternal caffeine with additional adjustment for maternal caffeine, (3) maternal caffeine with no further adjustment (i.e., the EWAS model), and (4) maternal caffeine with additional adjustment for paternal caffeine. Robust inferences were derived from models 2 and 4, which are adjusted for partner's caffeine consumption, as it reduces chances of confounding through assortative mating (when partners are not selected randomly but systematically based on specific characteristics, e.g., height or intelligence) (Lawlor, Wade, et al., 2019; Madley-Dowd et al., 2020). It has been suggested that caffeine consumption, as well as related behaviours – such as smoking and alcohol consumption – are subject to assortative mating (Madley-Dowd et al., 2020; Reynolds et al., 2006).

4.3.2.3.3 PRS analysis: Maternal polygenic risk score for caffeine intake during pregnancy (ALSPAC only)

To reduce chances of confounding and to further explore whether a causal effect of maternal caffeine consumption on offspring DNA methylation exists, I ran linear regressions in ALSPAC using a polygenic risk score (PRS) for caffeine as the exposure and the top maternal caffeine-associated CpG sites from the probe-level meta-analysis as the outcome. Genetic variants, in theory, should be free of confounding variables and thus provide less biased estimates (Davey Smith & Ebrahim, 2003). I derived the PRS by the weighted sum of risk-alleles that have been found to predict coffee and tea consumption in the general population (see section "polygenic risk score analysis" in Chapter 2 for further information) (Taylor, Davey Smith, et al., 2018; The Coffee and Caffeine Genetics Consortium et al., 2015). I adjusted all models, in addition to the covariates, also for offspring's caffeine PRS, which might be a confounder of the relationship of maternal caffeine PRS and offspring DNA methylation. Without adjustment for offspring PRS, a genetic effect of offspring PRS on offspring DNA methylation may be incorrectly attributed to the maternal caffeine PRS because of the shared genetic liability between mother and offspring (see Chapter 2 and 3 for more detail). Ideally, to avoid spurious results due to the genetic overlap between mother and offspring, results should be adjusted for both offspring and paternal PRS. Due to the scarcity of paternal genetic data in ALSPAC, I could only account for offspring PRS and results should be interpreted with caution as they might be affected by collider bias (Lawlor et al., 2017).

4.3.2.3.4 Functional analysis

To find out which gene pathways are linked to the CpG sites of the caffeine-associated DMRs, I ran a gene ontology analysis using the R package *missMethyl* (Phipson et al., 2015). Due to the unequal number of CpG sites per gene assessed on the 450k array, some genes are overrepresented and more likely to show up as enriched in gene set analyses (Geeleher et al., 2013). *missMethyl* accounts for this bias by considering the probability of a gene pathway being selected in accordance with the number of probes per gene on the array. I tested enrichment of gene ontology (GO) categories and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways.

4.4 Results

4.4.1 Sample characteristics

4.4.1.1 Maternal caffeine consumption during pregnancy

In all cohorts, most mothers (80-94%) consumed at least some caffeine during weeks 12-28 of gestation, with a mean of 85 mg/day over all cohorts, but with large variation within and between cohorts (weighted average SD = 82 mg/day, Table 4.1). Approximately 14% of mothers in the total sample consumed more caffeine than the 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 (Figure 4.1 and Table 4.1).

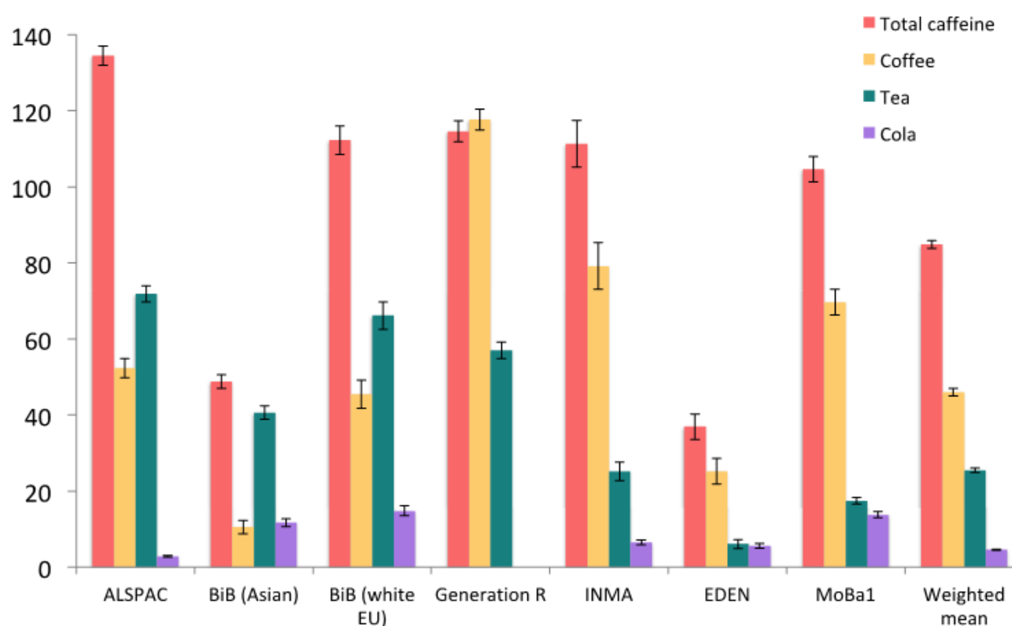


Figure 4.1 Visualisation of average maternal caffeine consumption during pregnancy across caffeine sources and cohorts. *Caffeine consumption reported in mg/day. Error bars represent standard errors. Asian = Asian ethnicity. White EU = White European ethnicity. Please note that I allowed for partially missing data when generating the total score and thus the individual sources of caffeine do not exactly add up to the total score.*

Table 4.1 Overview of maternal caffeine consumption during pregnancy in the individual cohorts

Cohort (N)	Weeks of gestation	M total caffeine (SD)	N users (%) > 200 mg of caffeine*	M coffee (SD)	M tea (SD)	M cola (SD)
ALSPAC (N = 729)	18	134.51 (94.43)	197 (27%)	52.37 (69.43)	71.84 (57.73)	2.85 (6.30)
BIB (Asian; N = 353)	26-28	48.82 (46.95)	5 (2%)	10.53 (33.10)	40.63 (33.49)	11.70 (19.98)
BIB (white EU; N = 306)	26-28	112.25 (105.13)	50 (19%)	45.48 (65.39)	66.14 (62.89)	14.82 (22.43)
Generation R (N = 798)	18-25	114.58 (95.67)	132 (20%)	117.69 (77.96)	57.01 (61.92)	Not available
INMA (N = 378)	12	111.30 (129.33)	26 (8%)	79.17 (119.74)	25.17 (46.24)	6.46 (12.11)
EDEN (N = 156)	24-28	36.93 (43.01)	1 (<1%)	25.21 (41.63)	6.09 (14.76)	5.62 (8.48)
MoBa1 1 (N = 999)	22	104.65 (106.27)	108 (11%)	69.67 (106.13)	17.46 (27.37)	13.78 (27.12)
Total or M** (N = 3,724)	-	84.86 (81.87)	519 (14%)	46.01 (64.96)	25.47 (34.95)	4.59 (10.61)

*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.*

4.4.1.2 Demographics

A general overview over the demographics of the individual cohorts can be found in Table 4.2. 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. There was no obvious or consistent difference in maternal age (except for MoBa1), level of education, maternal BMI, or offspring's gestational age, between mothers who did and did not consume caffeine during pregnancy (Table 4.3). However, 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%; Table 4.3). Furthermore, mothers who consumed caffeine were more likely to already have children (49%) compared to mothers who did not consume caffeine during pregnancy (33%; Table 4.3).

Table 4.2 Overview of maternal 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 (%)	M BMI	M gestational age
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 ethnicity; N = 353)	United Kingdom; Pakistani	EPIC	146 (41%)	28.21 (5.37)	9 (3%)	249 (70.5%)	25.75 (5.23)	26.62 (2.31)
BiB (**White EU ethnicity;** N = 306)	United Kingdom; Northern European	EPIC	125 (41%)	26.98 (6.15)	93 (30%)	159 (52%)	27.10 (6.48)	26.61 (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)	38.73 (1.55)

*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. Gestational age in BiB was assessed between 26-28 weeks of gestation. Parity = one or more previous pregnancies. EU = European.*

Table 4.3 Demographic variables of the cohorts stratified by caffeine user status

Cohort	Users							Non-users						
	N (%)	**M age (SD)**	**N any smoking preg.(%)**	**N high SEP(%)**	**N parity (%)**	**M BMI (SD)**	**M gest. age (SD)**	**N (%)**	**M age (SD)**	**N any smoking preg.(%)**	**N High SEP(%)**	**N parity (%)**	**M BMI (SD)**	**M gest. age(SD)**
ALSPAC	664 (91%)	29.71 (4.38)	72 (11%)	334 (50%)	348 (52%)	22.81 (3.61)	39.52 (1.52)	65 (9%)	30.68 (4.43)	5 (8%)	41 (61%)	33 (50.8%)	22.48 (3.79)	39.65 (1.43)
BiB (Asian)	284 (80%)	28.43 (5.11)	8 (3%)	117 (41%)	75 (26%)	25.77 (5.17)	26.61 (2.40)	69 (20%)	27.30 (6.32)	1 (1%)	29 (42%)	29 (42%)	25.70 (5.51)	26.64 (1.88)
BiB (White EU)	262 (86%)	26.88 (6.24)	86* (33%)	104 (40%)	139 (53%)	27.25 (6.57)	26.62 (1.94)	44 (14%)	27.55 (5.64)	7* (16%)	21 (48%)	20 (46%)	26.22 (5.96)	26.51 (1.48)
Generation R	650 (81%)	32.00 (4.20)	101* (16%)	448 (70%)	277* (42.6)	23.13 (3.70)	40.19 (1.51)	148 (19%)	31.65 (3.62)	8 *(5%)	106 (72%)	47* (32%)	22.73 (3.34)	40.21 (1.41)
INMA	326 (86%)	31.52 (4.02)	48 (15%)	241 (74%)	140 (42.9)	23.68 (4.19)	39.77 (1.36)	52 (14%)	31.73 (4.39)	5 (10%)	36 (69%)	21 (40%)	24.49 (5.77)	39.83 (1.60)
EDEN	144 (89%)	30.07 (5.04)	25 (17%)	82 (57%)	83 (57.6)	23.73 (4.76)	39.55 (1.36)	18 (11%)	30.78 (4.22)	1 (6%)	11 (61%)	12 (67%)	30.78 (4.22)	39.17 (1.10)
MoBa1	757 (76%)	30.19* (4.22)	241* (32%)	576 (76%)	461* (61%)	24.13 (4.32)	39.93 (1.60)	242 (24%)	29.12* (4.63)	46* (19%)	186 (77%)	49* (20%)	23.69 (3.67)	40.00 (1.40)
Total or M	3,087 (83%)	30.41 (4.43)	581 (20%)	1,902 (62%)	1,523 (49%)	23.64 (4.13)	38.57 (1.58)	638 (17%)	30.26 (4.42)	73 (11%)	430 (67%)	211 (33%)	23.64 (3.90)	38.21 (1.45)

*Note. M = Mean. SD = Standard deviation. In the Total row, average caffeine content was calculated by weighting by the inverse variance for each cohort. Mothers were categorised as non-user of caffeine if they indicated to consume zero mg/day of coffee, tea, or cola. * significant difference between users and non-users according to P-value < 0.05. High maternal socio-economic position (SEP): maternal education >= continued education after high school. Gestational age in BiB was assessed between 26-28 weeks of gestation. Parity = one or more previous pregnancies. EU = European. preg. = pregnancy. gest. = gestational age.*

4.4.2 Probe level meta-analysis

4.4.2.1 Quality control checks for cohort results

All correlation matrixes can be found in Appendix L, Figures L1 to L7. As expected, most of the models using the different caffeine phenotypes correlated moderately or highly. Only the model using caffeine from cola showed consistently very low correlations with the other model coefficients. Furthermore, the models using caffeine from tea and cola showed low and sometimes even negative correlations. This could be due to few mothers consuming caffeine from cola (Table 4.1) or indicate a differential effect of cola, tea, and coffee on DNA methylation, potentially through a different substance than caffeine or a different confounding structure. Lastly, the sex stratified models showed low correlations, most likely due to the small sample size in these models. Alternatively, this might indicate a differential effect of caffeine for male and female sex offspring. Visual inspection of the QQ-plots and Lambda values ranging from 0.99-1.07 indicated that most of the P-values of the models are randomly distributed (Appendix M, Figures A1 to M7). The precision plots can be found Appendix N, Figure N1). As expected, the cohorts with lower sample sizes had less precise estimates. MoBa1 showed most precise estimates throughout all models, except for the tea models, where Generation R showed the most precise estimates. Furthermore, the model using caffeine from cola as the exposure variable were found to be less precise in all cohorts except for MoBa1, most likely because of low consumption of caffeine from cola in the other cohorts. ALSPAC showed relatively low precision in the any vs. no caffeine and coffee models, most likely because of few mothers abstaining from caffeine during pregnancy and mothers consuming more caffeine from tea than coffee (Appendix N, Figure N1).

4.4.2.2 Association between maternal caffeine consumption and offspring cord blood DNA methylation

The overall EWAS meta-analysis results of the caffeine models can be found in Table 4.4. After adjusting for multiple testing, the meta-analysis revealed only evidence for one CpG site (cg19370043) being negatively associated with total maternal caffeine consumption (estimate = -2.18×10^{-05}; SE = 4.10×10^{-06}; FDR adjusted P-value = 0.048) and two CpG sites (cg12788467, cg14591243) with

caffeine consumed from cola in mg/day (Respectively: estimate = 5.22×10^{-05}, SE = 9.30×10^{-06}, FDR adjusted P-value = 0.007; estimate = 2.77×10^{-05}, SE = 5.00×10^{-06}, FDR adjusted P-value = 0.007; Table 4.5). The maternal caffeine-associated CpG site survived the leave-one-out analysis (Appendix O, Figure O1). Of the cola-associated CpG sites only one CpG site (cg12788467) survived the leave-one-cohort-out analysis (Appendix O, Figure O1 and O2). According to the genecards database (Stelzer et al., 2016), the gene PRRX1 that is annotated to the total caffeine-associated CpG site cg19370043, has been found to be associated with determining mesodermal muscle types by regulating muscle creatine kinase. The gene HNF1B that has been annotated for CpG site cg12788467 has been found to be associated with regulation of nephron development of the embryonic pancreas (genecard database) (Stelzer et al., 2016). The effect estimates of models adjusted and unadjusted for covariates showed high to moderate correlations (Spearman's r = 0.38-0.93), and correlation estimates were highest for coffee and lowest for cola (range correlation estimates: Coffee = 0.70-0.75; tea = 0.38-0.41; cola = 0.15-0.17; Appendix P, Figure P1). This could indicate that cola consumption in our sample was higher confounded than consumption of the other caffeinated beverages. Gestational age does not appear have an influence on the relationship between maternal caffeine consumption and offspring DNA methylation as adding gestational age as an additional confounder to the models did not change the effect estimates (Spearman's r = 0.99-1.00). See Appendix P, Figure P1 for the full correlation matrix of the meta-analysed models.

Table 4.4 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)	2	1 (50%)	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. There is a slight deviation in sample sizes between the results and descriptive information (N = 6 participants) because of removal of data in MoBa between 2018 (time of analysis) and 2020 (time of generating descriptive information) due to withdrawal of consent and/or non-Nordic/non-European ancestry.*

Table 4.5 CpG sites associated with maternal total caffeine consumption and caffeine consumption from cola with FDR-adjusted P-value < 0.05

Predictor	**CpG (gene)**	**Estimate (SE)**	**FDR adjusted P- value**
Total caffeine consumption (mg/day)	Chr1: cg19370043 (PRRX1)	-2.18×10^{-05} (4.10×10^{-06})	0.048
Caffeine consumption from cola (mg/day)	Chr17: cg12788467 (HNF1B)	5.22×10^{-05} (9.30×10^{-06})	0.007
	Chr3: cg14591243 (STAG1)*	2.77×10^{-05} (5.00×10^{-06})	0.007

*Note. "Estimate" can be interpreted as difference in offspring percentage methylation per one standard deviation increase in consumption of caffeine in mg/day, after adjustment for all covariates. * did not survive the Leave-one-out analysis.*

4.4.2.3 Beverage-specific effects

Figure 4.3 displays DNA methylation at the 3 CpG sites discovered in the probe-level analysis across the different caffeine models. If these associations were truly driven by caffeine exposure, I would expect to see similar associations of these CpG sites across the different beverage models. However, Figure 4.3 shows that the association between DNA methylation at cg19370043 and maternal total caffeine consumption (red) was mostly driven by coffee (yellow). There was no evidence for an association with DNA methylation at cg19370043 in the tea (green) or cola (purple) models. The effect estimate for the association between coffee and DNA methylation at cg19370043 was very similar to that for the association between total caffeine and DNA methylation at this site. The P-value was < 0.05, but the association 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 partially missing data). The two associations between DNA methylation at cg12788467 and cg14591243 appear to be specific to cola consumption (purple) rather than general caffeine consumption.

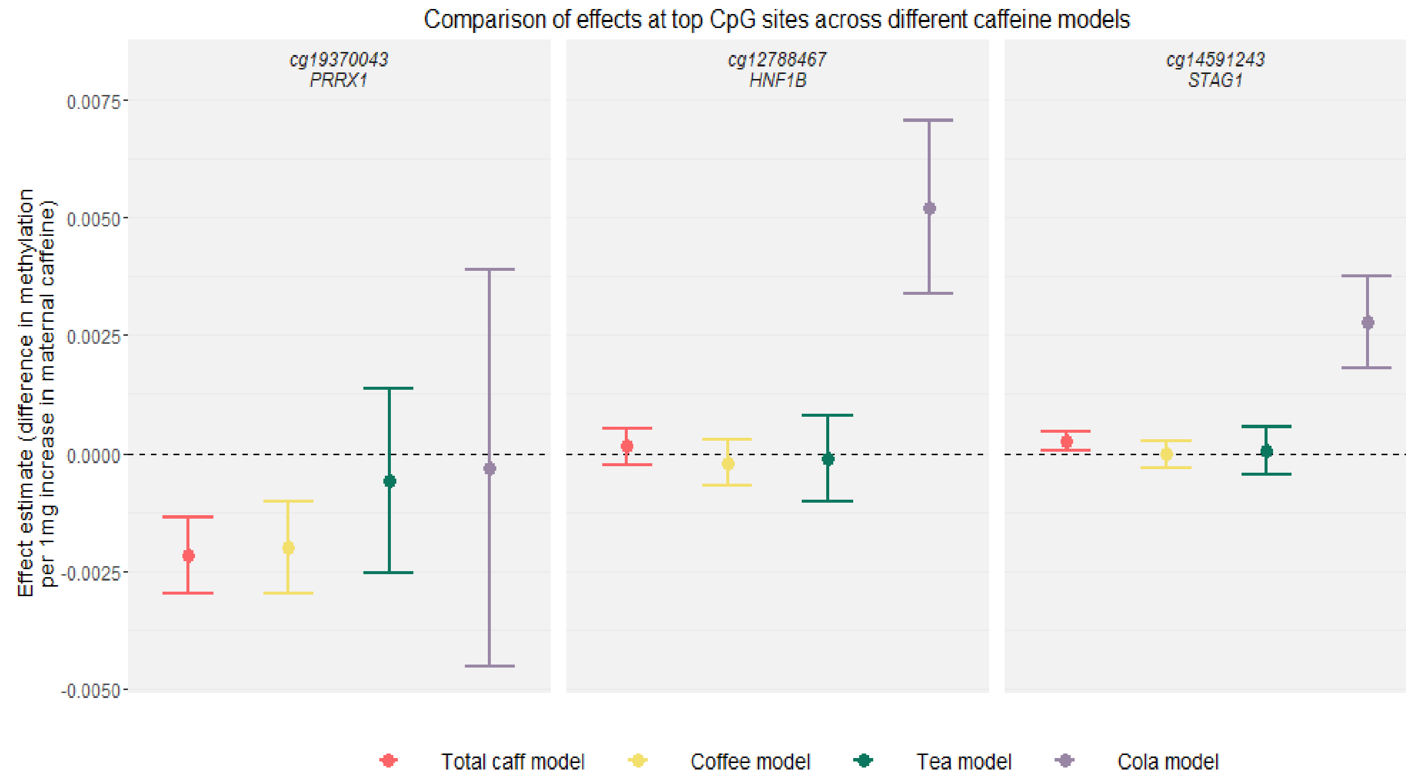


Figure 4.2 Effect size estimates at the top CpG sites found in the probe-level analysis. *Total caff model = total caffeine model. Error bars represent 95% Confidence Intervals.*

4.4.2.4 Negative control analysis (ALSPAC only): Associations between paternal caffeine intake and offspring DNA methylation (N = 619)

In ALSPAC, adjustment for the other parent's caffeine consumption did not show strong evidence for a change in magnitude of the total caffeine and cola associated CpG sites, indicating little evidence for an influence of shared parental confounders on the association with DNA methylation at this CpG site (Table 4.6 and Figure 4.3). The effect estimates of the total caffeine CpG site (cg19370043) was slightly larger in the maternal caffeine models than in the paternal caffeine models (Table 4.6 and Figure 4.3), providing some evidence for a somewhat stronger effect of maternal than paternal caffeine consumption during pregnancy at this CpG site (which theoretically is in line with an intrauterine effect, yet the differences of the estimates were very small). The effects at the cola associated CpG site, cg14591243, appeared to be of similar magnitude for paternal and maternal cola consumption in the negative control comparison in ALSPAC, indicating that the association at this CpG site that were observed in the meta-analysis, might be explained by common confounding factors of maternal and paternal caffeine consumption (e.g., SEP, diet, etc.) instead of actual caffeine exposure. The effect at the cola associated CpG site, cg12788467, did not change after adjusting the maternal caffeine model for paternal caffeine consumption. Yet, the effect was slightly attenuated after adjusting the paternal model for maternal caffeine consumption (Table 4.6 and Figure 4.3). This provides some evidence that the association with maternal cola consumption at this CpG site is not influenced by shared maternal and paternal confounding factors and that the effect is slightly stronger for maternal than paternal cola consumption. However, inference about an intrauterine effect based on this analysis is limited due to the small sample size, which is reflected by the wide confidence intervals, which are substantially overlapping for the maternal and paternal associations, especially in the analyses of the cola-associated CpG sites.

Table 4.6 Looking-up CpG sites found to be associated with maternal caffeine consumption in the results of associations between paternal caffeine consumption and offspring DNA methylation

Exposure	CpG (gene)	Estimate (SE)	P-value
Paternal self-reported caffeine consumption during pregnancy adjusted for maternal caffeine consumption			
Paternal total caffeine consumption adjusted for maternal total caffeine consumption	cg19370043 (PRRX1)	3.72×10^{-06} (1.31×10^{-05})	0.776
Paternal caffeine consumption from cola (mg/day) adjusted for maternal cola consumption	cg12788467 (HNF1B)	6.22×10^{-05}(1.13×10^{-04})	0.584
	cg14591243 (STAG1)	4.51×10^{-05} (5.68×10^{-05})	0.423
Maternal self-reported caffeine consumption during pregnancy adjusted for paternal caffeine consumption			
Maternal total caffeine consumption adjusted for paternal total caffeine consumption	cg19370043 (PRRX1)	-3.67×10^{-05} (2.10×10^{-05})	0.662
Maternal caffeine consumption from cola adjusted for paternal cola consumption	cg12788467 (HNF1B)	1.53×10^{-04} (8.46×10^{-05})	0.071
	cg14591243 (STAG1)	4.85×10^{-05} (4.24×10^{-05})	0.254
Maternal self-reported caffeine consumption during pregnancy unadjusted for paternal caffeine consumption			
Maternal total caffeine consumption	cg19370043 (PRRX1)	-3.48×10^{-05} (1.85×10^{-05})	0.061
Maternal caffeine consumption from cola	cg12788467 (HNF1B)	1.52×10^{-04} (7.60×10^{-05})	0.046
	cg14591243 (STAG1)	6.01×10^{-05} (3.75×10^{-05})	0.109

Note. caffeine consumption assessed in mg/day.

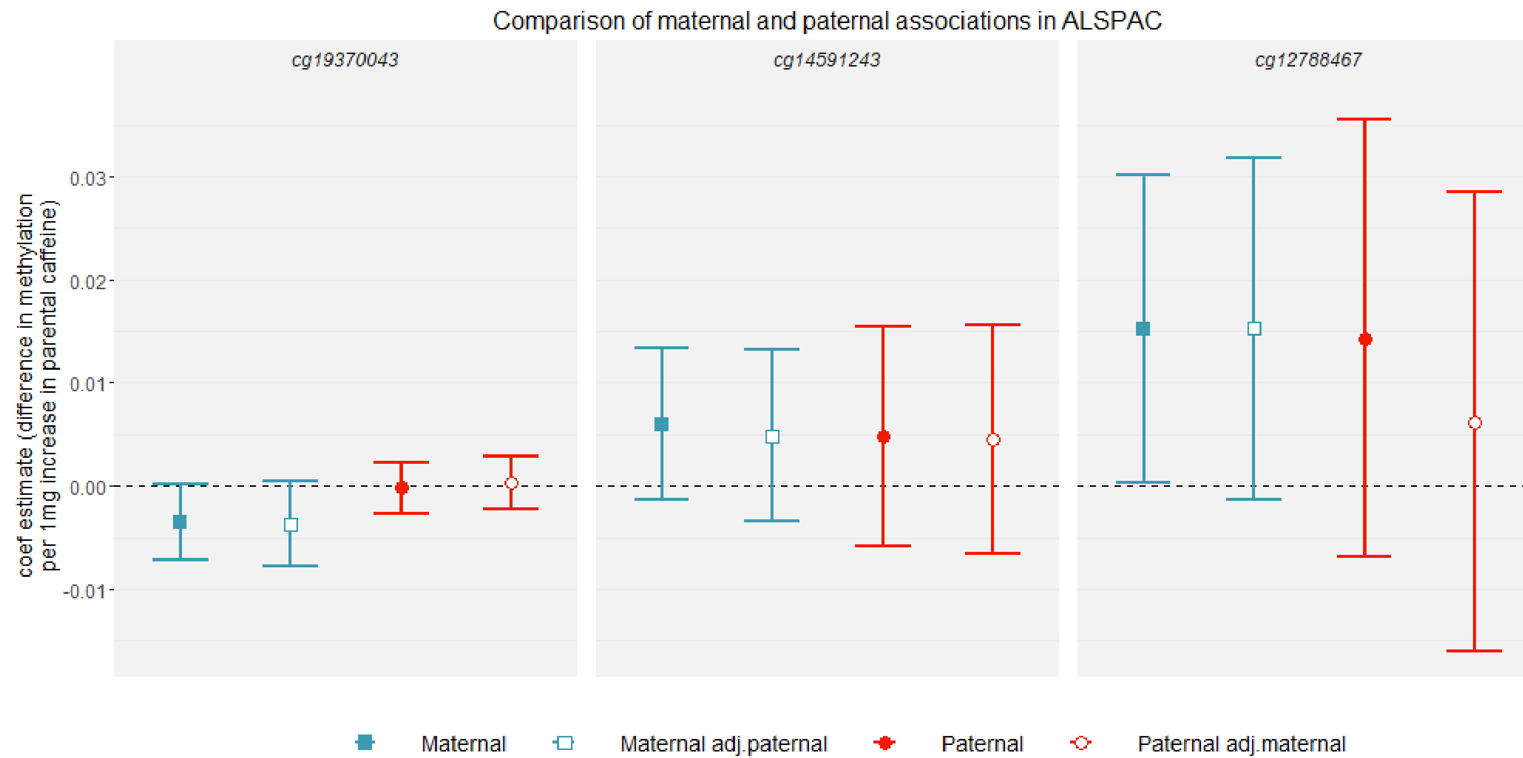


Figure 4.3 Comparison of maternal and paternal caffeine EWAS effect estimates at the top CpG sites of the meta-analysis results in ALSPAC. *Error bars represent 95% Confidence Intervals. Maternal = not adjusted for paternal caffeine consumption. Maternal adj. paternal = adjusted for paternal caffeine consumption. Paternal = not adjusted for maternal caffeine consumption. Paternal adj.maternal = adjusted for maternal caffeine consumption.*

Using all available CpG sites on the 450k array, results of the maternal models adjusted for paternal consumption showed no evidence for associations after multiple testing corrections (all FDR corrected P-values > 0.05). Interestingly though, the paternal models of any vs. no paternal caffeine consumption, showed evidence for many other CpG site associations, even with adjustment for maternal consumption (Table 4.7). However, given that the models with the other paternal caffeine phenotypes did not show evidence for associations, it seems more likely that due to the small sample size of paternal non-caffeine drinkers (N = 13), these results are false positives. Visual inspection of dot plots comparing DNA methylation values at the top 5 CpG sites of offspring whose fathers consumed any caffeine with fathers who did not consume any caffeine (Figure 4.4) did not indicate that these effects were driven by outliers (except for cg23063666, Figure 4.4 top left). However, as can be seen in the plots, differences were small and the variance within groups high, thus, no clear evidence for an effect of paternal caffeine consumption on offspring DNA methylation can be concluded from this data.

Table 4.7 Summary of results of paternal caffeine models in ALSPAC

Model*	CpGs with FDR-corrected P-value <0.05	Sample size	Genomic inflation factor (λ)**
Any vs. no caffeine (non-user = 13, user = 606)			
All offspring (minimally adjusted) *	40	619	1.23
All offspring (adjusted for covariates)	17	617	1.20
Female offspring (adjusted for covariates)	54	312	1.18
Male offspring (adjusted for covariates)	27	305	1.07
All offspring (adjusted for covariates and gestational age)	18	617	1.18
Caffeine in mg/day			
All offspring (minimally adjusted)*	0	619	1.00
All offspring (adjusted for covariates)	0	617	1.01
Female offspring (adjusted for covariates)	0	312	0.94
Male offspring (adjusted for covariates)	0	305	0.98
All offspring (adjusted for covariates and gestational age)	0	617	1.00
Caffeine from coffee			
All offspring (adjusted for covariates)	0	608	0.99
Caffeine from tea			
All offspring (adjusted for covariates)	0	562	0.96
Caffeine from cola			
All offspring (adjusted for covariates)	0	587	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 and the corresponding maternal caffeine consumption phenotype)*
*** 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.*

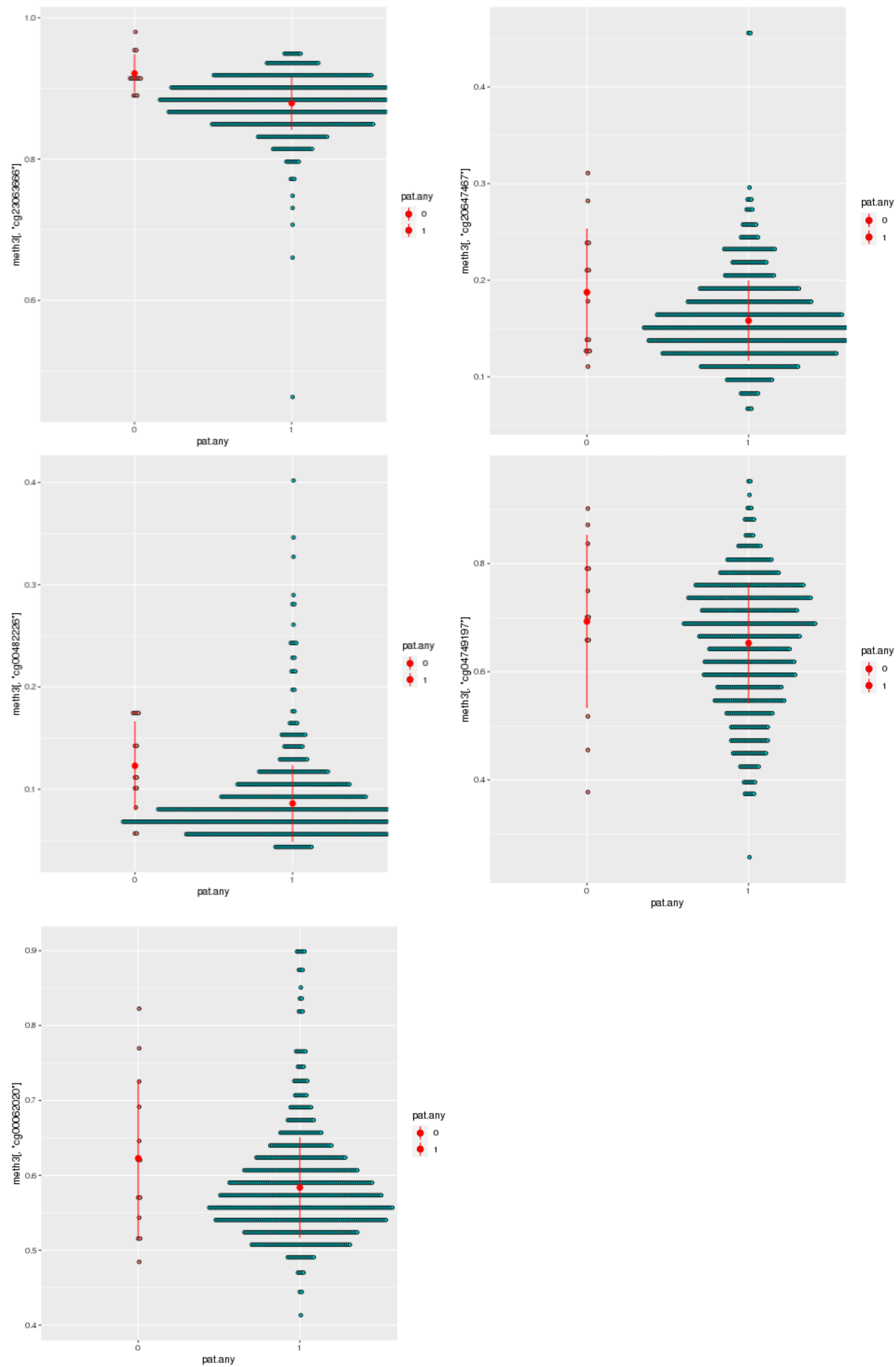


Figure 4.4 Dot plots comparing offspring DNA methylation at the top 5 CpG sites of the any paternal caffeine models in ALSPAC. *Models were adjusted for maternal caffeine consumption. 0 = non-user, 1 = caffeine user.*

4.4.2.5 PRS analysis (ALSPAC only): Associations between maternal caffeine PRS and offspring DNA methylation (N = 521)

The 3 maternal caffeine associated CpG sites from the probe-level analysis showed no evidence for association with the maternal PRS for caffeine (all P-values > 0.05) and even showed a different direction of effect for the cola-associated CpG sites (Table 4.8). Also, results from the analysis using all CpG sites on the array showed no evidence for associations between maternal caffeine PRS and offspring cord blood DNA methylation after correcting for multiple testing (FDR adjusted P-value > 0.05). Results of the six CpG sites with a P-value of less than 5×10^{-5} are available in Appendix Q, Table Q1.

Table 4.8 Results of the maternal caffeine PRS analysis in ALSPAC

Exposure	**CpG (gene)**	**Estimate (SE)**	**P-value**
Maternal caffeine PRS	cg19370043 (PRRX1)	-0.26 (0.23)	0.259
	cg12788467 (HNF1B;HNF1B)	-0.05 (0.03)	0.461
	cg14591243 (STAG1)	-0.02 (0.03)	0.608

4.4.3 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 2 and a maximum of 17 consecutive CpG sites ($P_{Bonferroni} < 0.05$; Figure 4.5, Appendix R, Table R1). 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}; $P_{Bonferroni} = 1.42 \times 10^{-10}$; Figure 4.5, Appendix R, Table R1). In the any vs. no maternal caffeine consumption model, DMRff detected evidence for 11 DMRs (Figure 4.6, Appendix R, Table R2). 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}; $P_{BF} = 1.928^{-06}$, Appendix R, Table R2). Only the COL9A3 gene on chromosome 20 was found to be in common between the DMRs detected in the total and any vs. no maternal caffeine models (Table 4.9).

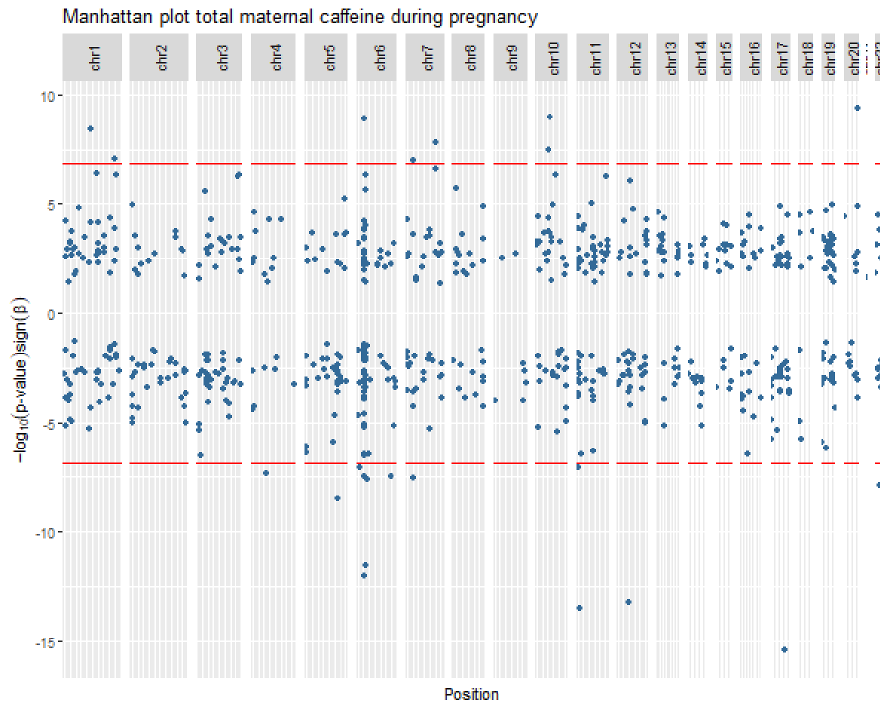


Figure 4.5 Candidate DMRs for total maternal caffeine consumption. *Each dot represents a candidate DMR with at least two CpG sites showing the same direction of effect. The red lines represent the Bonferroni corrected P-value threshold. Chr = Chromosome. Dots above the top line represent positive associations with offspring DNA methylation whereas dots below the lower line represent negative associations with offspring DNA methylation.*

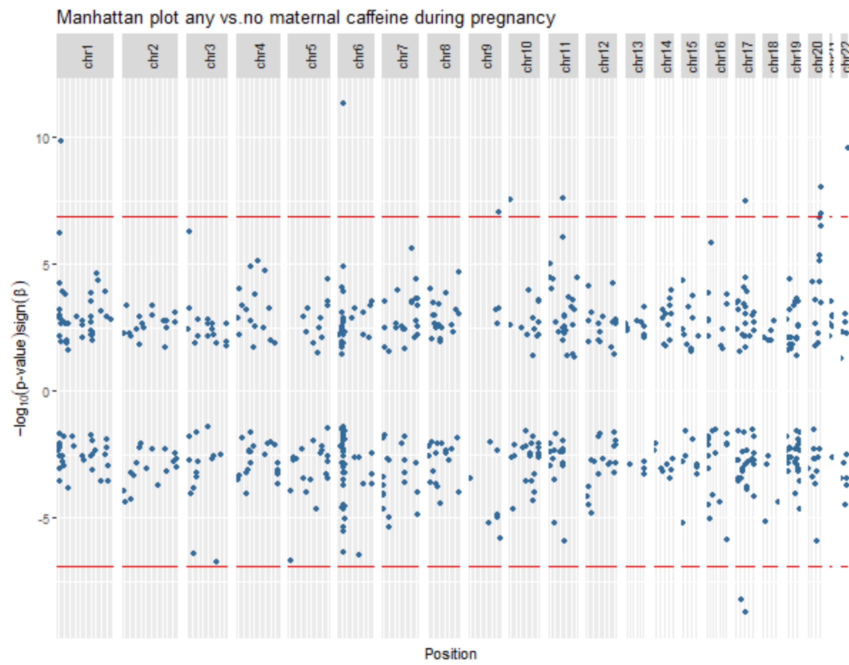


Figure 4.6 Candidate DMRs for any maternal caffeine consumption. *Each dot represents a candidate DMR with at least two CpG sites showing the same direction of effect. The red lines represent the Bonferroni corrected P-value threshold. Chr = Chromosome. Dots above the top line represent positive associations with offspring DNA methylation whereas dots below the lower line represent negative associations with offspring DNA methylation.*

DMR analyses from the individual sources of caffeine revealed 12 DMRs for caffeine consumed from coffee (Figure 4.7), 18 DMRs for caffeine from tea (Figure 4.8), and 14 DMRs for caffeine consumption from cola (Figure 4.9) during pregnancy. All results are presented in Appendix R, Table R1 to R7.

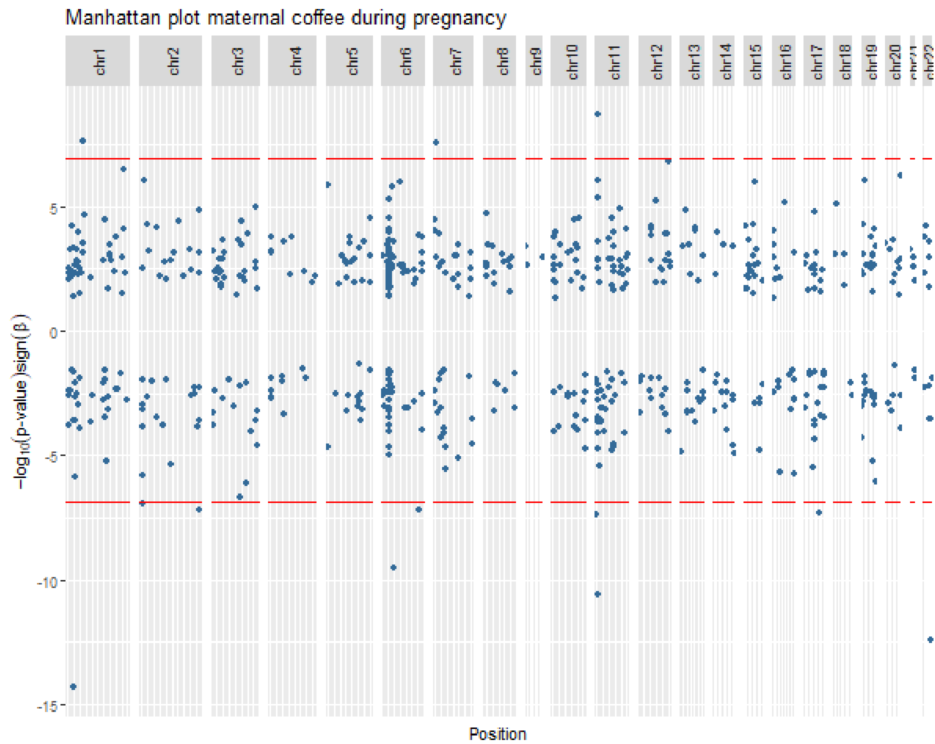


Figure 4.7 Candidate DMRs for maternal coffee consumption. *Each dot represents a candidate DMR with at least two CpG sites showing the same direction of effect. The red lines represent the Bonferroni corrected P-value threshold. Chr = Chromosome. Dots above the top line represent positive associations with offspring DNA methylation whereas dots below the lower line represent negative associations with offspring DNA methylation.*

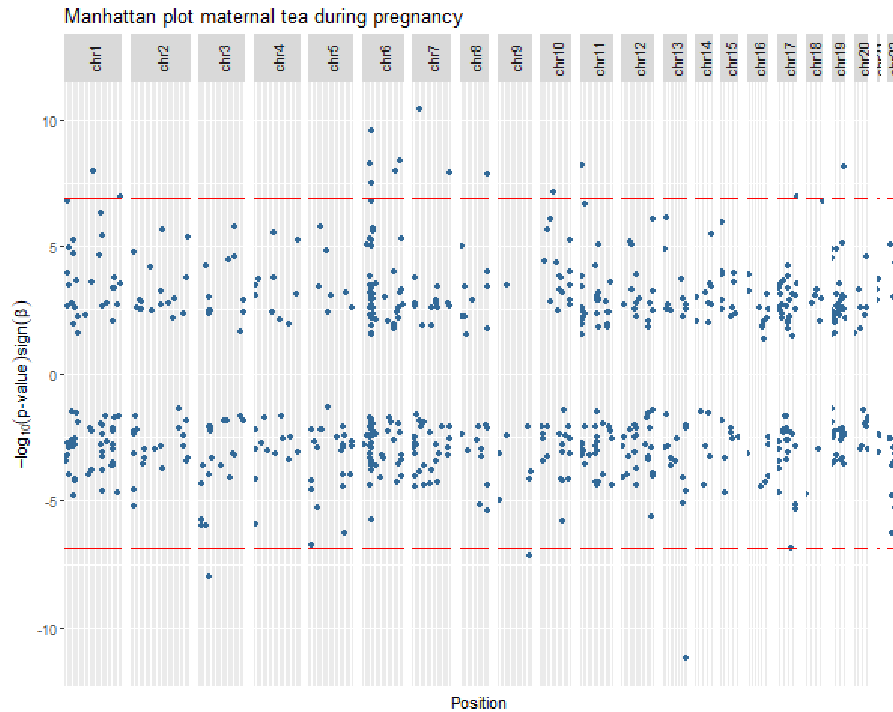


Figure 4.8 Candidate DMRs for maternal tea consumption. *Each dot represents a candidate DMR with at least two CpG sites showing the same direction of effect. The red lines represent the Bonferroni corrected P-value threshold. Chr = Chromosome. Dots above the top line represent positive associations with offspring DNA methylation whereas dots below the lower line represent negative associations with offspring DNA methylation.*

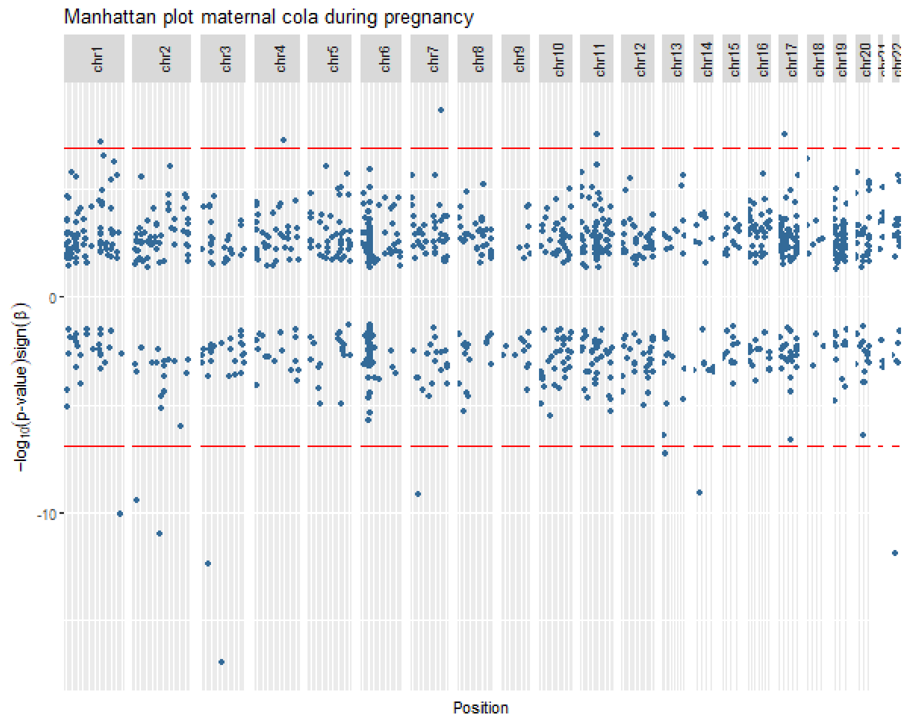


Figure 4.9 Candidate DMRs for maternal cola consumption. *Each dot represents a candidate DMR with at least two CpG sites showing the same direction of effect. The red lines represent the Bonferroni corrected P-value threshold. Chr = Chromosome. Dots above the top line represent positive associations with offspring DNA methylation whereas dots below the lower line represent negative associations with offspring DNA methylation.*

As shown in Table 4.9 there was little overlap between the CpG sites or genes annotated to the DMRs of the different caffeine models. For each combination of models, I calculated the percentage overlap of CpG sites by dividing the overlap of CpG sites between two models by the sum of the models unique CpG sites (e.g., percentage crossover any caffeine and total caffeine models: 7/(167 + 63 - 7) = 0.03 * 100 = 3%). Despite coffee being the most popular source of caffeine across cohorts, more crossovers were found between the total caffeine and tea models (crossover 11% and three crossover genes; *C17orf64, GABBR1; HOXA2*) than between the total caffeine and coffee models (crossover 7% and two crossover genes; C17orf64; B4GALNT4). Common genes and CpG sites annotated to the DMRs from the individual sources of caffeine were found between caffeine consumed from coffee and tea, which showed a 12% crossover of CpG sites (N crossover CpG sites = 19) that were annotated to the H19 gene on chromosome 11 and the C17orf64 gene on chromosome 17 (Table 4.9). However, in contrast to the DMRs associated with the C17orf64 gene, DMRs associated with the H19 gene showed a different direction of effect for maternal coffee and tea consumption (Appendix R, Table R3 and R4). No crossovers of genes and CpG sites of the cola associated DMRs were found with any of the other caffeine models (Table 4.9 and Appendix R, Tables R1 to R7).

Table 4.9 Crossover of CpG sites and genes of the DMRs of the different caffeine models

	CpGs total caffeine (genes)	**CpGs any vs. no caffeine (genes)**	**CpGs coffee (genes)**	**CpGs tea (genes)**	**CpGs cola (genes)**	**Total caffeine – Female sex (genes)**	**Total caffeine – Male sex (genes)**
CpGs total caffeine (genes)	167 CpGs (100%)	7 CpGs (COL9A3), crossover = 3%	17 CpGs (C17orf64; B4GALNT4), crossover = 7%	26 CpGs (C17orf64 GABBR1; HOXA2), crossover = 11%	0 CpGs (0 genes)	6 CpGs (B4GALNT4), crossover = 3%	0 CpGs (0 genes)
-	**CpGs any vs. no caffeine (genes)**	63 CpGs (100%)	0 CpGs (0 genes)	0 CpGs (0 genes)	0 CpGs (0 genes)	0 CpGs (0 genes)	0 CpGs (0 genes)
-	-	**CpGs coffee (genes)**	86 CpGs (100%)	19 CpGs (H19; C17orf64), crossover = 12%	0 CpGs (0 genes)	4 CpGs (B4GALNT4), crossover = 3%	0 CpGs (0 genes)
-	-	-	**CpGs tea (genes)**	92 CpGs (100%)	0 CpGs (0 genes)	0 CpGs (0 genes)	0 CpGs (0 genes)
-	-	-	-	**CpGs cola (genes)**	55 CpGs (100%)	0 CpGs (0 genes)	0 CpGs (0 genes)
-	-	-	-	-	**Total caffeine – Female sex (genes)**	55 CpGs (100%)	0 CpGs (0 genes)
-	-	-	-	-	-	**Total caffeine – Male sex (genes)**	85 CpGs (100%)

Note. Cells highlighted in blue represent the total number of CpG sites (CpGs) that were contained across the DMRs of each caffeine model. Darker shading of blue = more overlap. Crossover = percentage of crossover CpG sites between the different models (N CpGs crossover / (N CpGs model 1 + N CpGs model 2 – N CpGs crossover).

The analyses from the sex-stratified models showed evidence for 12 DMRs in cord blood in female sex offspring (Figure 4.10; Appendix R, Table R6) and 18 DMRs in male sex offspring (Figure 4.11 and Appendix R, Table R7) being associated with total maternal caffeine consumption. Some crossover was found between CpG sites and genes of the female sex stratified caffeine DMRs and the unstratified total caffeine model (crossover = 3%; six common CpG sites with the total caffeine model and four common CpG sites with the coffee model; Table 4.9). However, no crossover was found between the genes and CpGs sites annotated to the male sex stratified total caffeine model and any other caffeine model (Table 4.9). This could indicate a sex-specific effect of maternal caffeine consumption during pregnancy on offspring DNA methylation. Alternatively, this could also represent male and female DNA methylation differences independent of maternal caffeine consumption. As has been mentioned in Chapter 2 (section 2.3.2.4), differences in DNA methylation between male and female sex can still be observed when restricting analyses to autosomes and removing probes cross-reactive with sex chromosomes (Yousefi, Huen, Davé, et al., 2015).

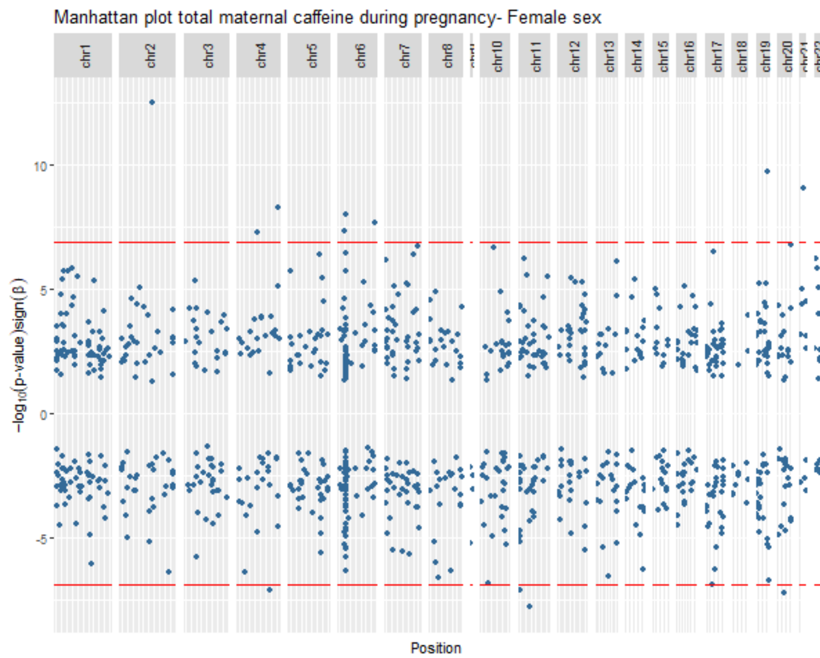


Figure 4.10 Candidate DMRs for female sex stratified models of total maternal caffeine consumption. *Each dot represents a candidate DMR with at least two CpG sites showing the same direction of effect The red lines represent the Bonferroni corrected P-value threshold. Chr = Chromosome. Dots above the top line represent positive associations with offspring DNA methylation whereas dots below the lower line represent negative associations with offspring DNA methylation.*

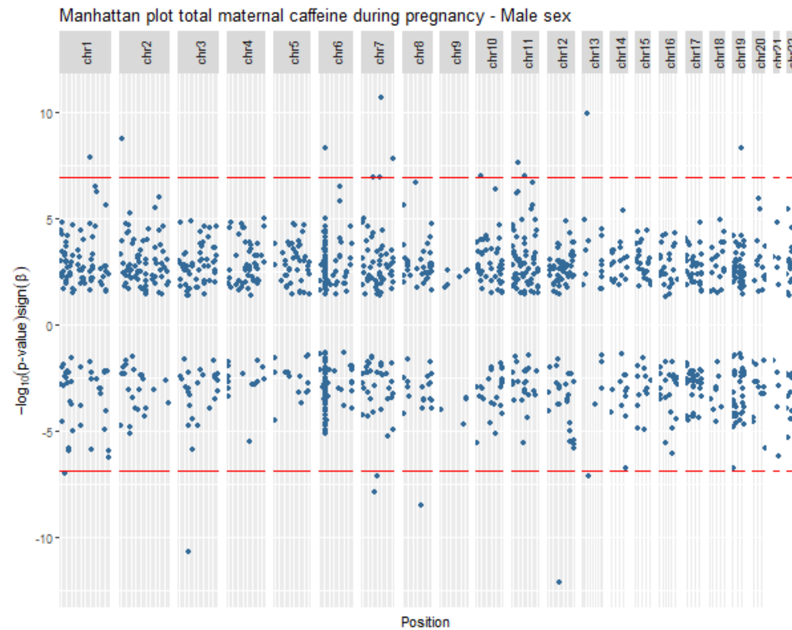


Figure 4.11 Candidate DMRs for male sex stratified models of total maternal caffeine consumption. *Each dot represents a candidate DMR with at least two CpG sites showing the same direction of effect. The red lines represent the Bonferroni corrected P-value threshold. Chr = Chromosome. Dots above the top line represent positive associations with offspring DNA methylation whereas dots below the lower line represent negative associations with offspring DNA methylation.*

4.4.3.1 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 Appendix S, Table S1.

4.4.4 Analysis of cell proportions

The estimated cell proportions meta-analysis revealed that none of the caffeine phenotypes were associated with the proportions of cells in offspring cord blood (e.g., P-values of maternal caffeine meta-analysis: BCELL = 0.988; CD14 = 0.917; CD4T = 0.058; CD8T = 0.848; GRAN = 0.323; NK = 0.187).

4.5 Discussion

4.5.1 Summary & interpretation of findings

In this study, I investigated the effects of maternal caffeine consumption during pregnancy on offspring cord blood DNA methylation, using data from six international birth cohorts from the PACE consortium. Different methods were used to triangulate the evidence: Associations were investigated between offspring DNA methylation and (1) maternal self-reported caffeine consumption during pregnancy across six birth cohorts (EWAS meta-analysis), (2) paternal self-reported caffeine consumption in ALSPAC (negative control analysis), and (3) maternal caffeine PRS in ALSPAC. Furthermore, the EWAS meta-analysis compared caffeine consumption during pregnancy across contexts (and caffeinated drinks) and thus reduced the potential for cultural confounding in the association between maternal caffeine consumption during pregnancy and offspring DNA methylation.

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. Both these analyses were conducted for total caffeine consumption during pregnancy, as well as for caffeine consumption from coffee, tea, and cola consumption separately. Inferences about causal associations were attempted by first, testing the total caffeine and cola-associated CpG sites from the probe level analysis in a negative control and PRS analysis and second, analysing the CpG sites of the maternal caffeine associated DMRs for their biological function.

4.5.1.1 Findings from the probe-level analysis

The probe level analysis indicated that differences in DNA methylation at three CpG sites were associated with maternal caffeine phenotypes (one with maternal total caffeine consumption and two with maternal caffeine consumption from cola) of which all 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, if the recommended limit of caffeine consumption during pregnancy would be doubled from 200 mg/day of

caffeine to 400 mg/day caffeine a day, this would only be associated with a 0.004% decrease in DNA methylation at cg19370043 (coefficient multiplied by 200; based on the assumption that intrauterine caffeine exposure has a linear effect on offspring DNA methylation). For the caffeinated cola results, drinking one extra cup of cola per day would be associated with a 0.017% increase in DNA methylation at cg12788467 (coefficient multiplied by the 20 mg caffeine contained in one cup of cola). These effect sizes are in line with the small effect sizes found in the EWAS of caffeine consumption on own DNA methylation (Karabegović et al., 2020), where an additional cup of coffee was associated with a 0.002% decrease in peripheral blood DNA-methylation at *AHRR*, which would be equivalent to a 0.007% decrease in DNA methylation per 200 milligrams of caffeine (0.002% / 57 mg of caffeine per cup of coffee x 200). The effect sizes of caffeine found in this study and the study of Karabegović and colleagues appear to be much smaller than the effect of smoking, where sustained smoking during pregnancy was associated with changes of up to ~7% decrease in offspring cord-blood DNA methylation at the *AHRR* gene (Joubert et al., 2016).

4.5.1.1.1 Relevance for offspring's mental health development

I found no evidence to link the caffeine associated CpG sites to any mental health phenotypes. According to the EWAS catalog, none of the three maternal caffeine associated CpG sites have previously been reported to be associated with neurological functions or mental health phenotypes. Similarly, according to the GWAS catalog, none of the genes mapped to the CpG sites has been reported to be associated with any traits related to mental health.

4.5.1.2 Findings from the differentially methylated regions (DMR) analysis

In the regional analyses, DMRff identified 12-22 DMRs for each of the caffeine models. No clear cluster between the functions of genes associated with maternal caffeine DMRs could be found. Surprisingly, only one gene mapped to DMRs of the total and any vs. no caffeine models overlapped (COL9A3), which is most likely explained by analyses using continuous instead of binary exposure variables being higher powered and suffering less of residual confounding (Fedorov et al., 2009; Royston et al., 2006). Even though the regional analyses revealed several potential DMRs for each of the different caffeine models, there were few overlapping DMRs between models, which indicates that the

associations of the DMRs are not driven by actual intrauterine exposure of caffeine but rather other confounding structures of the different sources of caffeine. Caffeine consumption is a complex phenotype, even more so during pregnancy (e.g., due to caffeine metabolism changes). Besides caffeine, each of the included caffeinate beverages contains numerous other chemical compounds that might exert an effect on maternal and offspring health (Ludwig et al., 2014). Whereas the coffee and tea models both showed DMRs with the same direction of effect of the *C17orf64* gene on chromosome 17, no overlap was found with genes associated with the cola associated DMRs (Table 4.9). The biological function of *C17orf64* is not very clear yet, but previous research reported it to be negatively associated with a DMR associated with maternal obesity in female (but not male) offspring in cord blood (Martin et al., 2019). Even though models were adjusted for maternal BMI, associations between coffee and tea consumption and CpG sites at gene *C17orf64* might be due to residual confounding. Surprisingly, the DMRs of the male and female stratified models did not show any overlapping genes and, together with the larger effect estimates in the stratified compared to non-stratified total caffeine model, might indicate different effects of maternal caffeine consumption during pregnancy for female and male offspring (Table 4.9). Lack of overlap between the cola associated DMRs with the total caffeine associated DMRs (Table 4.9) further supports the assumption that the associations of maternal cola consumption on offspring DNA methylation might be explained by a different ingredient than caffeine or a different underlying confounding structure. This is in line with epidemiological evidence (N = 64,189), that found associations between maternal soft drink consumption during pregnancy and offspring hyperactivity, regardless of whether soft drinks contained caffeine, which indicates that associations are likely driven by other ingredients than caffeine (Berglundh et al., 2020). This is further supported by a study, which found that only maternal consumption of cola, but not consumption of coffee or tea, during pregnancy was associated with symptoms of inattention and hyperactivity in offspring at 18 months (N = 25,343) (Bekkhus et al., 2010).

4.5.1.2.1 Relevance for offspring's mental health development

Both total maternal caffeine consumption and maternal tea consumption during pregnancy showed a DMR mapped to the *GABBR1* gene (Appendix R, Table R1

and R4). According to the GWAS catalog, this gene has been associated with various mental health problems including schizophrenia, neuroticism, autism spectrum disorder, and depression. It is worth noting that, it has also been linked to smoking and nicotine dependence in the GWAS catalog, as well as with smoking during pregnancy in the EWAS catalog, and therefore might also represent residual confounding with maternal smoking during pregnancy. Yet, it is interesting that it did not come-up in the coffee DMR results (Appendix R, TableR3) despite the high correlation between smoking and coffee consumption. This could indicate that the association between caffeinated tea and *GABBR1* might be due to a different ingredient than caffeine or a different confounding structure of tea in this sample.

4.5.1.3 Attempts to infer causality

I made several attempts to investigate whether intrauterine caffeine exposure might have a causal effect on offspring DNA methylation. First, beverage-specific effects for total caffeine, coffee, tea, and cola were inspected across the DMR and probe-level results. Overall 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). Investigating the effects of the CpG sites found in the probe-level analysis across the different caffeine models further supported beverage-specific effects (Figure 4.2).

Second, I conducted a negative control analysis in ALSPAC, which used paternal instead of maternal caffeine consumption during pregnancy as the exposure. Including paternal caffeine consumption as a covariate in the maternal caffeine models did not show clear evidence to change the results. This indicates that the association are not strongly influenced by common confounding factors of maternal and paternal caffeine consumption. Whereas there was limited evidence for a difference in effects between maternal and paternal caffeine consumption at the cola-associated CpG sites (cg12788467 and cg14591243), there was some evidence for a stronger maternal than paternal effect at the total caffeine associated CpG site (cg19370043). This could provide some evidence for an intrauterine effect of caffeine on DNA methylation at cg19370043; yet again, differences in effect estimates were small and thus are unlikely to have a strong

effect on offspring outcomes. The hypothesis free exploration using all probes on the 450k array showed DNA methylation at many probes to be associated with any vs. no paternal caffeine consumption during pregnancy. However, considering the limited sample size (ALSPAC fathers only) with very few fathers abstaining from caffeine during pregnancy (N = 15), this analysis is highly underpowered, with a high risk for false positives. Yet, it is possible that paternal caffeine consumption is influencing offspring DNA methylation, for instance through the germ line (Crean & Bonduriansky, 2014), and this hypothesis deserves further investigation in larger, independent samples.

Last, I conducted a PRS analysis in ALSPAC using a caffeine PRS as a proxy for maternal caffeine consumption. Genetic variants for caffeine should in theory be free of confounding (Davey Smith et al., 2007; Davey Smith & Ebrahim, 2003) and thus give less confounded results than the analysis using self-reported caffeine. There was no evidence that the three CpG sites from the probe-level analysis, using self-reported maternal caffeine consumption as the exposure, were associated with the maternal PRS for caffeine in ALSPAC. Interestingly, however, whereas the total caffeine-associated CpG site showed the same direction of effect as in the maternal self-reported caffeine models, the cola-associated CpG sites showed a different direction of effect. This is in line with results of Chapter 3, in which the genetic variants of coffee consumption were associated with maternal consumption of coffee and tea outside and during pregnancy but not with cola consumption (Schellhas, Haan et al., 2021; Taylor, Davey Smith, et al., 2018). The maternal caffeine PRS analysis using all probes on the 450k array did not show any associations with offspring cord blood DNA methylation, which could either provide some evidence that there is no effect of maternal caffeine consumption on offspring cord blood DNA methylation or due to the limited sample size, was underpowered to detect small effects, especially for a genetic analysis.

4.5.2 Strengths

This was the first large, international EWAS meta-analysis investigating the effect of caffeine from coffee, tea, and cola during pregnancy. A major strength of this study is the thorough investigation of the effects of caffeine consumption during pregnancy through various methods with slightly different biases (probe-level,

regional, negative control, and PRS analysis). All methods revealing few associations with small effect sizes, increases confidence that low to moderate amounts of caffeine during pregnancy are unlikely to have large effects on offspring cord blood DNA methylation. Further, even though the study did not account for all potential sources of caffeine (e.g., caffeine from energy drinks, chocolate, green tea, medication, etc.), the consideration of the effects of other common sources of caffeine besides coffee is important to reduce bias. Consumption of the different sources of caffeine might be slightly differentially socially patterned, allowing capturing a larger spectrum of the caffeine consuming population. For instance, British and non-Western Ethnicities consume more commonly caffeinated tea than coffee (The Coffee and Caffeine Genetics Consortium et al., 2015; Treur et al., 2016). Further, research found that the main source of caffeine might change during pregnancy, with even habitual coffee drinkers preferring caffeinated tea over coffee during pregnancy (Chen et al., 2014; Lawson et al., 2004). Lastly, 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. Another strength is the use of two sensitivity analyses to explore whether any of the associations might reflect a causal effect of caffeine.

4.5.3 Limitations

The findings should be considered in the light of the following limitations. Measurement error might have occurred for several reasons. First, caffeine assessment in the meta-analysis relied on self-report, which has been found to be underestimated by non-pregnant men and women (Schreiber et al., 1988) and might even be more strongly underestimated during pregnancy because of social stigma (see Chapter 1). Second, the formula used to transform maternal caffeine consumption from cups/day to mg/day assumed relatively low amounts of caffeine per caffeinated drink in comparison to transformations in other studies (Treur et al., 2016; van Dam et al., 2020). Even though this might have led to a systematic underestimation of the effect size of caffeine for each of the beverages, the proportions between the individual caffeinated beverages was comparable to

that used in previous research (our study: 1 cup of coffee = 57 mg caffeine, 1 cup of tea = 27, 1 can of cola = 30, Ratio = 55%:26%:19%; e.g., Treur et al; 2016: 1 cup of coffee = 75 mg caffeine, 1 cup of tea = 40 mg caffeine, 1 can of cola = 33 mg caffeine: Ratio = 75%:40%:33%). Third, the questions used to assess caffeine consumption in the cohorts were very basic and did not differentiate between different cup sizes, types of coffee (e.g., brewed coffee, espresso, instant coffee, etc.), and brewing times, and therefore only allowed for rough estimations of maternal caffeine consumption (van Dam et al., 2020b). Fourth, the study relied on prospective cohort data, which is likely to suffer from selection bias, with more advantaged families tending to enrol in the study (Fraser et al., 2013). Fifth, results might not be generalizable to effects of maternal caffeine consumption at other trimesters of pregnancy or higher dosages of caffeine. In this study, maternal caffeine consumption was only investigated during the second trimester of pregnancy without controlling for pregnancy symptoms such as nausea and vomiting, which likely influence maternal caffeine consumption and might proxy for a healthy course of pregnancy (Lawson et al., 2004; Wikoff et al., 2017). Sixth, effects of caffeine exposure during pregnancy on offspring cord blood DNA methylation was only assessed at regions available on the 450k array, which only covers around 2% of CpG sites of the entire epigenome (Lövkvist et al., 2016). Thus, differentially methylated regions not covered by the array have been missed out in this study. Seventh, I 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 I cannot rule out that maternal caffeine during pregnancy might be influencing DNA methylation differentially in other tissue types (Walton et al., 2019). Last, this study indicates that, if maternal caffeine consumption influences cord-blood DNA methylation, the effect is likely to be small. Although efforts were made to maximise the sample sizes by including more than one cohort, even larger sample sizes may be required to detect small effects of prenatal caffeine exposure on offspring DNA methylation.

4.5.4 Future research

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 plasma concentrations of the caffeine metabolite paraxanthine (Boylan et al., 2008). Also, future research should further explore the different confounding structures of various caffeinated beverages across different cultures. Further, it would be important to consider individual differences in the maternal metabolism of caffeine that might influence the 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 (Cornelis et al., 2016) and/or consider pre-pregnancy caffeine consumption to account for differences in the tolerance to effects of caffeine during pregnancy (van Dam et al., 2020). Last, more assessments of prenatal paternal caffeine consumption are needed. This would enable the conduction of higher-powered negative control analyses to investigate intrauterine effects, as well as investigating the effects of paternal caffeine consumption prior to pregnancy and its effect on offspring DNA methylation.

4.5.5 Conclusion

In conclusion, results of this 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. Results of this study show little converging evidence between the associations of the different sources of caffeine, thus it seems more likely that associations between maternal caffeine consumption during pregnancy and offspring DNA methylation are explained by other factors than caffeine exposure, such as smoking or sugar content (which might explain the cola associations). By investigating caffeine consumption across different contexts (with slightly different confounding structures), chances of confounding bias should be reduced, however, as mothers across cohorts consumed relatively little caffeine during pregnancy, results might be different in higher caffeine consuming pregnant populations.

Chapter 5 – The development of childhood internalising problems: A meta-analysis of epigenome-wide-association-studies

5.1 Chapter overview

In this chapter, I explore the relationship between child DNA methylation (at birth and childhood) and internalising problems (at the age of 3 and 7) through a hypothesis-free EWAS meta-analysis. In addition to the EWAS analysis, a candidate-gene analysis is conducted testing DNA methylation at genes that have previously been found to be associated with phenotypes related to internalising problems (adverse early life experiences, adult anxiety and depression). The results of the EWAS analysis of this chapter are part of the third step of the meet-in-the-middle approach (see Chapter 1), which investigates the association between prenatal smoking and caffeine associated DNA methylations changes and offspring internalising problems (intermediate biomarker-disease association). I conducted parts of this chapter's analyses during secondments at the Erasmus Medical Center Rotterdam (EMCR; Generation R analysis) and the Norwegian Institute of Public Health (MoBa analysis). Due to time restrictions and data policy regulations that do not allow for remote access to the Generation R data, a PhD student from the EMCR, Mannan Lou, ran some of my scripts for me and sent me the Generation R summary results.

5.2 Introduction

Over the past 20 years the prevalence of internalising problems, including depression and anxiety disorders, has tremendously increased amongst adolescents and young adults (Sadler et al., 2018; Twenge et al., 2019). In the UK, around 8% of young people between the age of 5 to 19 suffer from internalising disorders (Sadler et al., 2018). Anxiety disorders are among the most common mental health problems in children (Creswell et al., 2014; Ramsawh et al., 2011), but are still rarely diagnosed before adolescence, with a peak in diagnosis during young adulthood (American Psychiatric Association, 2013). There is some indication that many anxiety and depressive disorders already onset during childhood. For instance, in a nationally representative sample of more than 10,000 US adolescents, anxiety disorders were reported to already been evident

during childhood, with a median age of onset at the age of 6 (Merikangas et al., 2010). Furthermore, a prospective study in New Zealand (N = 1,037) found that about 50% of adults with depression and anxiety disorders at the age of 26, experienced psychological difficulties already between 11-15 years of age, with 12% of cases reporting depression and 18% reporting anxiety symptoms before 15 years of age (Kim-Cohen et al., 2003). Early detection of anxiety and depression is particularly important as early onset has been found to be associated with a more severe course, higher chances for chronicity and recurrence, as well as a higher risk for suicide attempts, than late onset (Ramsawh et al., 2011; Weissman et al., 1999). Anxiety and depression commonly co-occur, and there is evidence that they follow a bi-directional relationship, with onset of anxiety disorders being a risk factor for developing depression and vice versa (Jacobson & Newman, 2017). Also, prevalence rates of comorbid anxiety and depression are high during childhood (Cummings et al., 2014; Garber & Weersing, 2010), with approximately 15% to 25% of young people diagnosed with depression having a comorbid anxiety disorder, and 10% to 15% of young people with an anxiety disorder suffering from comorbid depression (Cummings et al., 2014). The bi-directional relationship between anxiety and depression, which is already evident during childhood, stresses that future research should aim to detect early manifestations of anxiety or depressive symptoms, so that onset can be prevented (Jacobson & Newman, 2017).

The aetiology of anxiety and depression is determined by genetic and environmental influences. Twin studies suggest that anxiety and depression in childhood are moderately heritable, with about 40% of variance being explained by genetic influences (Polderman et al., 2015), and that there is a shared genetic liability for anxiety and depression (Eley, 1999; Gottschalk & Domschke, 2017). In contrast to twin studies, SNP-based heritability from recent GWAS of anxiety and depression report lower heritability estimates of 26% to 31% for anxiety (Purves et al., 2020), and only 9% for depression (Howard et al., 2019). In addition to genetic risk factors, early life experiences may shape the risk for the development of anxiety and depression later in life (Burt, 2009). Adverse intrauterine experiences, as well as early post-natal experiences, have been found

to be observationally associated with an increased risk for childhood internalising problems (Duko et al., 2020; LeMoult et al., 2020).

A molecular mechanism that could account for the interplay between genetic and environmental influences on childhood internalising symptoms is DNA methylation. DNA methylation at birth has been proposed to be a valuable proxy for the quality of the intrauterine environment (Felix et al., 2018; Miguel et al., 2019; Teh et al., 2014) and thus may help to shed light on the contribution of adverse intrauterine experiences to the development of offspring internalising problems later in life. In this chapter, I am going to investigate whether differences in DNA methylation at birth and childhood are associated with early manifestations of depression and anxiety, in the form of internalising problems, in childhood.

5.2.1 Epigenetic studies of offspring anxiety and depressive symptoms

5.2.1.1 Candidate gene studies of anxiety and depression

Most of the studies investigating the relationship between DNA methylation and anxiety and depressive symptoms in children have investigated differences in DNA methylation at candidate genes (Barker, Walton, & Cecil, 2018; Weder et al., 2014). A common theory for how adverse experiences become biologically embedded pre- and postnatally is through influencing the stress-response regulated by the hypothalamic-pituitary-adrenal axis (HPA-axis) (Cao-Lei et al., 2017; Miguel et al., 2019) and therefore many of these candidate gene studies have focussed on genes that are involved in the regulation of the HPA-axis (Kumsta, 2019; McGowan et al., 2009a; Murgatroyd & Spengler, 2011; Weaver et al., 2004). In the following section, I review the evidence for the most commonly investigated candidate genes, and briefly discuss how their function might contribute towards the development of internalising problems.

5.2.1.1.1 Nuclear receptor subfamily 3 group C member 1 (*NR3C1*) gene

In healthy individuals, stress induces activity of the HPA-axis to release cortisol, a glucocorticoid hormone, into the blood. Once a certain amount of cortisol is reached, a negative feedback loop becomes activated by cortisol reducing its own release through binding to the glucocorticoid receptor, that suppresses the activity of the HPA-axis and therewith stops the release of cortisol (Miller &

O'Callaghan, 2002). Research found that in individuals suffering from depression, the activation of the negative feedback loop is disturbed, probably due to lack of sensitivity of the glucocorticoid receptor, causing inflated levels of cortisol in the blood (Pariante & Lightman, 2008). Offspring can either be exposed to elevated cortisol levels prenatally, through maternal stress during pregnancy leading to an increased level of maternal glucocorticoids, which cross the placenta barrier, or postnatally, by offspring's own secretion of glucocorticoids. There is evidence that both, increased prenatal and postnatal glucocorticoid exposure, is associated with dysregulation of offspring's HPA-axis and brain development (Duthie & Reynolds, 2013; Lupien et al., 2009). Prenatal stress, particularly maternal anxiety and depression, was found to be associated with differences in offspring's neuronal development (Miguel et al., 2019) and an inflated stress response after stressful events (Gutteling et al., 2005).

A well-known epigenetic study by Weaver and colleagues (Weaver et al., 2004) found support for DNA methylation at the glucocorticoid receptor mediating the effect of early life adversity on the stress-response. Specifically, they investigated the effects of a high- vs. low-nurturing postnatal environment and found that, in contrast to rat pups raised in a high nurturing environment, pups raised in a low nurturing environment showed an increased fear response and elevated activity of the HPA-axis in response to stress (Weaver et al., 2004). The effects of high vs. low maternal care were mediated through epigenetic changes in offspring at the exon 1_7 of the promoter site of the Glucocorticoid receptor gene called *Nr3c1*. Following the findings of the study of Weaver and colleagues (Weaver et al., 2004), many other candidate gene studies were conducted to investigate epigenetic changes at the glucocorticoid receptor gene in response to prenatal and postnatal exposure to stress. A systematic review from 2016 reported that at least 13 animal and 27 human studies have investigated this topic (Turecki & Meaney, 2016). According to the systematic review, 89% of these candidate gene studies showed congruent findings for an association between increased DNA methylation at the 1_F exon variant of *NR3C1* gene (the human counterpart to exon 1_7 of *Nr3c1* in rats) and different types of early adversity. Two of these studies investigated cord blood DNA methylation: One reporting hypermethylation at 1_F of the *NR3C1* gene in cord blood of 33 children exposed to maternal depression,

compared to DNA methylation levels of 36 children not exposed to prenatal depression (Oberlander et al., 2008). The other study found evidence for pregnancy related anxiety being associated with increased offspring cord blood DNA methylation at promoter regions of the *NR3C1* gene amongst 83 mothers and their offspring, with the strongest evidence at exon 1_F (Hompes et al., 2013). Findings from these two *NR3C1* candidate gene studies in cord blood were partly, but not completely, overlapping (Hompes et al., 2013).

In addition to candidate gene studies of cord blood, *NR3C1* methylation assessed in peripheral blood was reported to be positively associated with internalising problems amongst 468 adolescents exposed to postnatal stressful life events, however, no DNA methylation differences at *NR3C1* were found in adolescents exposed to perinatal stress (van der Knaap et al., 2014). Two studies investigated DNA methylation near promoters of the *NR3C1* gene in brain tissue of suicide victims and only found hypermethylation in individuals who experienced early child abuse but not in suicide victims without experience of child abuse (Labonte et al., 2012; McGowan et al., 2009). Together with the other candidate gene studies, this suggests that changes in DNA methylation at *NR3C1* are linked to an elevated stress-response and may be particularly influenced by childhood adversity (Labonte et al., 2012; McGowan et al., 2009; Weder et al., 2014).

Besides *NR3C1*, other genes that have been targeted in DNA methylation studies investigating environmental exposures on children's stress-response, neurodevelopment, and mental health problems are (Barker, Walton, & Cecil, 2018; Weder et al., 2014): The serotonin transporter gene (*SLC6A4)* (Brummelte et al., 2017; Oberlander et al., 2010; Roberts et al., 2014), Brain-derived neurotropic factor (*BDNF)*(Fuchikami et al., 2011) and the FKBP Prolyl Isomerase 5 gene (*FKBP5)* (Klengel & Binder, 2015).

5.2.2 Epigenome-wide association studies of anxiety and depression

Candidate gene studies are unlikely to capture the full complexity of the gene-environment interplay that contributes towards the development of internalising disorders (Barker, Walton, & Cecil, 2018). Yet only two EWAS have investigated associations between DNA methylation and childhood internalising problems. A small EWAS investigated the association between DNA methylation in saliva of

190 children (aged 5 to 14 years) discordant for maltreatment. While DNA methylation at three CpG sites was associated with depressive symptoms in the overall sample, no differences between DNA methylation levels of maltreated and non-maltreated children were observed (Weder et al., 2014). Genes corresponding to these CpG sites were *ID3*, *NMDA*, *GRIN1,* and *TPPP*, which are also involved in the stress-response and neurological development. Another small EWAS of 18 monozygotic twins discordant for adolescent depression found increased buccal cell DNA methylation at one CpG site, annotated to the *STK32C* gene, to be associated with depression, which was replicated in post-mortem brain tissue of veterans discordant for depression (Dempster et al., 2014). No overlap in results was observed between these two EWAS, neither between the CpGs sites nor their annotated genes. While DNA methylation sites at *NR3C1* did not survive multiple testing corrections in the two EWAS, some sites showed at least nominal significance in the EWAS of Dempster and colleagues (Dempster et al., 2014) and survived Bonferroni correction in a candidate gene follow-up in Weder and colleagues' study (Weder et al., 2014).

5.2.3 Summary

In summary, DNA methylation may be a potential biological marker for internalising problems in childhood and help to improve the understanding and identification of symptoms that are commonly not diagnosed before adolescence. Candidate studies have consistently found altered DNA methylation at *NR3C1* in offspring exposed to some form of prenatal or postnatal early life stress. Yet, in EWAS studies, changes in methylation at *NR3C1* tend to not survive adjustment for multiple testing. Rather, results of EWAS studies indicate that DNA methylation at other genes might be more relevant for the development of anxiety and depression. This highlights the importance of not just focussing on selected genes, but to take on a hypothesis-generating epigenome-wide approach to discover the contribution of novel genes to the development of internalising problems (Barker, Walton, & Cecil, 2018). Studies that have taken an epigenome-wide approach for investigating internalising symptoms in offspring were of small size, only investigated DNA methylation in childhood/adolescence (i.e., after postnatal influences on DNA methylation have taken place), and solely focussed on depressive but not anxiety symptoms in childhood. The aim of this study was

to increase understanding for the contribution of prenatal and postnatal environmental exposures to the development of internalising problems in childhood by investigating associations between cord blood and peripheral blood DNA methylation (in childhood) with internalising symptoms at ages 3 and 7, using data from multiple prospective European studies which collectively provide a large sample size (maximum N = 3,011).

5.3 Methods

To investigate critical windows in development when children are particularly susceptible to biologically embed adverse environmental experiences that are associated with internalising problems in childhood, two main analyses were conducted. The first analysis was set out to investigate the relationship between cord blood DNA methylation and childhood internalising problems, where associations might indicate the contribution of *prenatal* experience*s* to internalising problems in childhood. The cord blood analysis involved a prospective EWAS using offspring cord blood DNA methylation as the exposure and internalising problems in children at the age of 3 or 7 years as the outcome (from here on referred to as "*cord blood analysis age 3"* and "*cord blood analysis age 7"*). Second, to investigate *postnatal* effects, results from the cord blood analyses were compared with results from a cross-sectional analysis using childhood peripheral blood DNA methylation at the age of 7 and internalising problems at the age of 7 (from here on referred to *childhood peripheral blood analysis age 7*).

5.3.1 Meta-analysis of epigenome-wide association studies

5.3.1.1 Participating cohorts

The prenatal EWAS meta-analysis included three independent birth cohorts that had data on offspring cord blood DNA methylation and internalising problems around the age of 3 available (total N = 3,011): ALSPAC (N = 739); Generation R (N = 793), MoBa1 (N = 998), and Moba2 (N = 481). As the MoBa DNA methylation subsamples did not have data on child internalising problems after the age of 3 available, and did not have DNA methylation data beyond birth, some meta-analyses contained data from ALSPAC and Generation R only. These were: A cord blood meta-analysis at age 7 (ALSPAC: N = 709; Generation R: N = 892;

total N = 1,601) and a peripheral blood meta-analysis at age 7 (ALSPAC: N = 747; Generation R: N = 374; total N = 1,121). An overview of the sample sizes of the individual cohorts for the cord blood and peripheral-blood analyses is presented in the Flowcharts (Figure 5.1 and Figure 5.2). I excluded multiple pregnancies (e.g., twins) and siblings so that each mother was only represented once in the datasets. For more details about the individual cohorts please refer to Chapter 2.

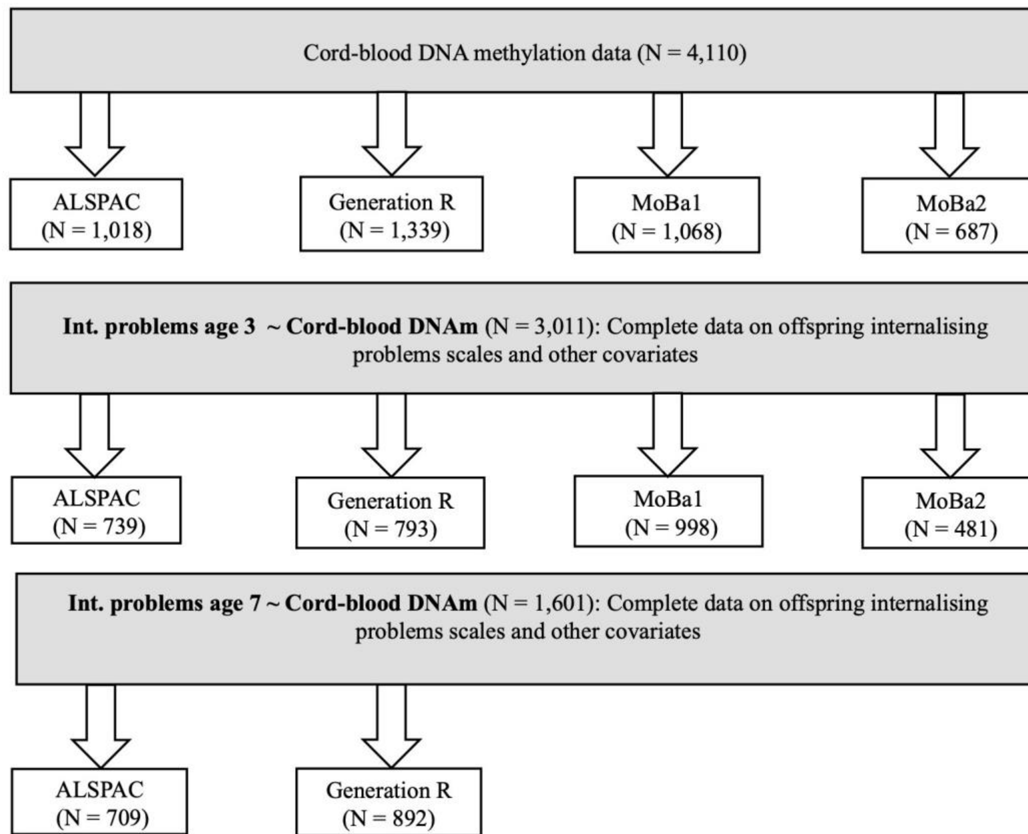


Figure 5.1 Flow chart of the sample sizes of the cord blood analyses. *Int. problems = Internalising problems.*

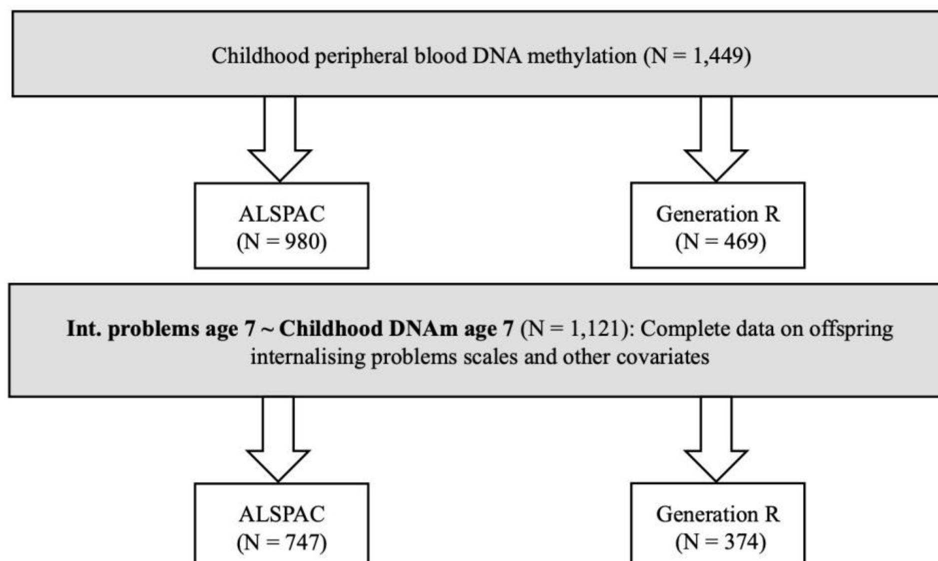


Figure 5.2 Flow chart of the sample size of the cross-sectional analysis. *Int. problems = Internalising problems.*

5.3.1.2 Measures

5.3.1.2.1 DNA methylation

Cord blood and peripheral blood DNA methylation was assessed through normalised beta values. Cohorts assessed 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). Probes on SNPs, probes that cross-hybridized according to Chen and colleagues (Chen et al., 2013), and probes on the sex chromosomes were excluded.

5.3.1.2.2 Childhood internalising problems

Child behaviour checklist (CBCL-1½-5). In Generation R and MoBa, child internalising problems at the age of 3 were assessed using the internalising subscale of the CBCL (Achenbach & Rescorla, 2000). The internalising subscale combines scores on the Anxious/Depressed, Withdrawn-Depressed, Emotional Reactivity and Somatic Complaints syndrome scales. Mothers reported on their children's internalising problems over the last two months, on a 3-point Likert scale, ranging from 0 = "Not True" to 2 = "Very True or Often True", with higher scores reflecting more internalising problems. Whereas Generation R used the full internalising scale (N items = 36), MoBa1 and MoBa2 used a short scale, which contained a subset of 9-items of the full internalising scale (further information on the MoBa subscale can be found here https://www.fhi.no/globalassets/dokumenterfiler/studier/den-norske-mor-far-og-barn--undersokelsenmoba/instrumentdokumentasjon/instrument-documentation-q6.pdf).

Strengths and Difficulties Questionnaire (SDQ). In ALSPAC, internalising problems at the age of 3 were assessed using the emotional symptoms subscale of the SDQ (Goodman, 1997). The emotional symptoms scale incorporates 5 items that are rated by mothers on a 3-point Likert scale (0 = Not true, 1 = Somewhat true, 2 = Certainly true), with higher scores reflecting more internalising problems. Mothers were asked to report on their child's behaviours over the past six months.

Comparability of the questionnaires. Previous research from the UK (Goodman & Scott, 1999), Germany, and Austria (Klasen, 2000) has shown that the internalising problems subscale of the CBCL and emotional symptoms subscale of the SDQ correlate moderately to highly (r = 0.69 - 0.74) in 4 to 10-year-old children and can equally well distinguish between clinical and control populations. The comparability of the scales is further supported by studies from the Netherlands (van Widenfelt et al., 2003) and Finland (Koskelainen & Kaljonen, 2000), which found the CBCL and SDQ to be comparable in samples of older children aged 11-16 years.

5.3.1.2.3 Covariates

I included the following state-of-the-art covariates for EWAS analyses in the models to reduce biological and technical variation (see Chapter 2 for more detail).

Offspring age. Due to known associations between age and variation in DNA methylation (Teschendorff, West, et al., 2013), I adjusted the cord blood analyses for gestational age in days and the meta-analysis using peripheral blood in childhood for offspring's age at DNA methylation assessment.

Technical batch. I accounted for DNA batch effects by adding 20 surrogate variables.

Cell proportions. As explained in Chapter 2, cellular heterogeneity is an important source of variation in DNA methylation data that can exaggerate or mask true differences in methylation levels. I estimated cell proportions using the Houseman method (Houseman et al., 2012) with a cord blood reference panel for the analyses using cord blood (Gervin et al., 2016) and a peripheral blood reference panel for the childhood peripheral blood meta-analysis (Reinius et al., 2012). I then included estimated cell counts in models as covariates.

I included the following covariates in the analyses to reduce variability in the outcome or if there was indication of confounding (variable might be causing both, DNA methylation and outcome). Therefore, I included covariates if there was indication in the literature that they were associated either with internalising problems (to reduce variability in the outcome) or correlated with both, DNA

methylation and internalising problems. A description of the variables of each cohort can be found in Appendix T, T1 to T3.

Maternal education. An ordinal variable of maternal education, with higher scores representing a higher level of education, was used as a proxy for family socioeconomic position (SEP). Family SEP has been found to be associated with higher risk for offspring mental health problems (Barker et al., 2012; Goodman et al., 2011) and is commonly adjusted for in EWAS (Joubert et al., 2016; Sharp et al., 2021).

Maternal age (years). Maternal age was found to be associated with both offspring mental health problems (Carslake et al., 2017), and DNA methylation (Markunas et al., 2016).

Maternal smoking during pregnancy. Maternal smoking during pregnancy was included as an ordinal variable (0 = no smoking, 1 = giving up smoking early in pregnancy, 2 = smoking throughout pregnancy) and included because of the strong influence on DNA methylation (Joubert et al., 2016) and indication for association with offspring internalising problems (Ashford et al., 2008b; Moylan et al., 2015).

Maternal anxiety and depression during pregnancy. Further, I separately adjusted for maternal anxiety and depressive symptoms during pregnancy, which should also be a proxy of postnatal depression (Matijasevich et al., 2015). Maternal depression, both during and outside of pregnancy, has been associated with an increased risk for offspring mental health problems (Ahun et al., 2018; Barker et al., 2012; Goodman et al., 2011; Matijasevich et al., 2015). Further, maternal anxiety and depression may partly explain variation in offspring DNA methylation because of the shared genetic propensity for anxiety and depression between mother and offspring, which may explain part of the variation of offspring's DNA methylation levels (Jones et al., 2018; Miguel et al., 2019). In ALSPAC, frequency of maternal anxiety and depressive symptoms during pregnancy were assessed with the Edinburgh Postnatal Depression scale (10 items rated on 4-point Likert scale ranging from 0 = "Not at all" to 3 = "yes, most of the time") (Cox et al., 1987) in the second trimester of pregnancy. In Generation R, the anxiety and depressive symptoms subscale of the Brief Symptom Inventory

(12 items, rated on a 5-point likert scale ranging from 0 = "Not at all" to 4 = "extremely") (Derogatis & Melisaratos, 1983), assessed during the second trimester of pregnancy, were used. In MoBa1 and MoBa2, five selected items of the Hopkin's symptom checklist (4-point Likert scale ranging from 0 = "Not bothered" to 3 = "Very bothered") (Derrogatis et al., 1973), assessed in the third trimester of pregnancy, were used (Appendix T for further detail).

5.3.2 Statistical analyses

5.3.2.1 Cohort-specific statistical analyses

Before running the EWAS, I removed probes that fell outside 3 times the inter-quartile range (IQR: 25th to 75th percentile) (Tukey, 1977) from the methylation data. I generated surrogate variables using the R-package SVA (Leek et al., 2012) and estimated cell proportions using the Houseman method (Houseman et al., 2012). For each model, I ran a linear regression model at each CpG site using the R-package Limma (Ritchie et al., 2015), with offspring DNA methylation as the exposure and internalising problems in childhood as the outcome. I z-standardised internalising problems scores and maternal depression and anxiety symptoms during pregnancy before running the regression models to ensure compatibility of the scales used in the different cohorts. The cord blood and childhood peripheral blood analyses included five models each: A minimally adjusted model only adjusted for estimated cell proportions, a covariate adjusted model including all covariates (except for maternal anxiety and depression symptoms during pregnancy), two models stratified for sex, and a final model that was additionally to the covariates, adjusted for maternal anxiety and depression symptoms during pregnancy.

5.3.2.2 Probe-level analysis

5.3.2.2.1 Quality control checks for cohort results

Prior to meta-analysing summary results from each cohort, I conducted quality checks to ensure that there were no problems with the data. First, I generated correlation matrices of the beta coefficients for each of the models. Second, I plotted and inspected the distributions of the P-values through QQ-plots, and calculated Lambda values. Further, I generated precision plots by plotting

1/median standard error against the square root of the sample size of each cohort (see Chapter 2 for more detail).

5.3.2.3 Differentially methylated regions (DMR analysis)

In addition to the probe level analysis, I conducted a regional analysis in each cohort. More details about this method can be found in Chapter 2. I ran DMR analyses using the dmrff.pre function from the dmrff R-package (Suderman et al., 2018). I annotate probes to the human reference genome version 19, build 37h using the annotation data available from the R-package meffil (Suderman et al., 2019).

5.3.2.4 Meta-analysis

5.3.2.4.1 Probe-level meta-analysis

I meta-analysed results with fixed effect estimates, weighted by the inverse of the variance using METAL (Willer et al., 2010). I adjusted P-values using a 5% FDR (see Chapter 2) and considered associations with an adjusted P-value below 0.05 as evidence for statistical significance (Benjamini & Hochberg, 1995).

5.3.2.4.2 Quality control checks for meta-analysed results

I inspected P-value distributions of the meta-analysed results through creating QQ-plots and corresponding Lambda values. Also, for the meta-analysed models, I generated correlation matrices of effect estimates. Furthermore, I performed a leave-one-cohort-out analysis using the R-package metafor (Viechtbauer, 2010) to determine the impact of each individual cohort on the meta-results (see Chapter 2).

5.3.2.4.3 Differentially methylated regions (DMR) meta-analysis

I meta-analysed DMR summary statistics from each cohort using an inverse-variance weighted fixed effects approach within the dmrff.meta function (Suderman et al., 2018). I defined a DMR as a region with at least two CpG sites with the same direction of effect and a Bonferroni adjusted P-value of less than 0.05 ($P_{Bonferroni}$ = original P-value x number of tests that were run).

5.3.3 Candidate-gene wide analysis

To test whether DNA methylation at four commonly tested candidate genes show association with internalising symptoms in this sample, I looked-up results of

CpG sites annotated to *NR3C1, SCL6A4, BDNF,* and *FKBP5* in the probe-level meta-EWAS results. To assess evidence for an association between these CpG sites and internalising problems I adjusted the P-value threshold according to the Bonferroni method, where the number of tests was defined as the number of CpG sites annotated to a given candidate gene.

5.4 Results

5.4.1 Sample characteristics

5.4.1.1 Maternal characteristics

Demographic information about mothers that were included in the meta-analysis can be found in Table 5.1. Overall, offspring and maternal characteristics appear to be very similar across the different cohorts. The sample of the cord blood meta-analysis at age 3 contained more mothers with higher education levels (69%) than mothers with lower education levels, and this pattern was consistent across the individual cohorts, except for ALSPAC, where mothers were approximately equally distributed between high and low education groups (high education = 50%, Table 5.1). On average, 11% of mothers in the sample smoked in the third trimester or throughout pregnancy (Table 5.1). For the cord blood and peripheral blood meta-analyses using internalising problems at the age of 7 as the outcome (not including MoBa), the sample became slightly more equally distributed in terms of maternal education, yet still slightly more mothers had a high school degree (and/or further education; ~58%) than no High School degree (Table 5.1). The distribution of mothers who smoked during pregnancy was comparable to the sample in the analysis using internalising problems at the age of 3 as the outcome (11% to 12%, Table 5.1). Also, maternal age at birth was consistent across the cohorts and assessment time points, with mothers being on average 31 years of age at the birth of the study child (range: 29.75 to 32.31 years; Table 5.1).

Table 5.1 Maternal characteristics

Cohort	**High maternal SEP (N (%))**	**Maternal age (Mean, SD)**	**Maternal smoking (N (%))**	**Maternal anxiety/depression during pregnancy (Mean (SD))***
Internalising problems age 3 ~ Cord blood DNAm				
ALSPAC (N = 739)	372 (50%)	29.75 (4.32)	79 (11%)	0.63 (0.46)
Generation R (N = 793)	560 (71%)	32.06 (3.99)	83 (11%)	0.13 (0.22)
Moba 1 (N = 998)	838 (84%)	29.97 (4.30)	146 (15%)	0.23 (0.35)
Moba 2 (N = 481)	391 (81%)	29.90 (4.29)	43 (9%)	0.23 (0.34)
Total or mean (N = 3,011)	2,161 (72%)	30.52 (4.21)	351 (12%)	-
Internalising problems age 7 ~ Cord blood DNAm				
ALSPAC (N = 709)	366 (52%)	29.92 (4.35)	74 (10 %)	0.90 (0.47)
Generation R (N = 892)	578 (65%)	31.78 (4.21)	103 (12%)	0.13 (0.26)
Total or mean (N = 1,601)	944 (60%)	30.99 (4.27)	177 (11%)	-
Internalising problems age 7 ~ Peripheral blood DNAm age 7				
ALSPAC (N = 747)	384 (51.4%)	29.89 (4.33)	80 (11%)	0.60 (0.49)
Generation R (N = 374)	264 (70.6%)	32.31 (3.90)	40 (11%)	0.11 (0.22)
Total or mean (N = 1,121)	648 (57.8%)	30.81 (4.17)	120 (11%)	-

*Note. In the Total row, means were weighted by the inverse variance for each cohort. Smoking = Smoking late/throughout pregnancy. High maternal socio-economic position (SEP): maternal education >= high school diploma. * As each cohort used different scales, it was not possible to generate a total sample mean.*

As different scales assessed maternal anxiety and depressive symptoms across the cohorts, no weighted mean could be calculated. The item means of the individual scales can be found in Table 5.1. Across cohorts and time-points, most mothers did not show symptoms of anxiety and depression during pregnancy. Mothers in ALSPAC had the highest frequency of experiencing anxiety or depressive symptoms, reporting on average to experience symptoms “not very often” (Mean = 0.63, SD = 0.46; Table 5.1) during the past seven days. In Generation R, mothers reported on average to experience any anxiety or depressive symptoms “not at all” (Mean = 0.13, SD = 0.22) during the past seven days, and in MoBa1

and MoBa2, mothers reported on average to “not have been bothered” with any anxiety or depressive symptoms in the past two-weeks (Mean = 0.23, SD = 0.63-0.64; Table 5.1). This patterned remained consistent for analyses of internalising symptoms at the age of 7 (Table 5.1).

5.4.1.2 Offspring characteristics

At the assessment of internalising symptoms at the age of 3, the sample contained children with an average age of approximately 3.5 years (weighted mean in months = 41.67, SD = 1.18; Table 5.2). At the assessment of internalising problems at the age of 7, children within the sample were on average 6.7 years of age (weighted mean in months = 80.93, SD = 1.28; Table 5.2). Overall, the samples contained an equal number of male and female sex offspring, except for MoBa2, which contained slightly more male than female sex offspring (44% female offspring, Table 5.2). Across cohorts, female and male sex children showed similar levels of internalising problems (Table 5.2). At the age of 3, children showed almost no internalising problems, with mothers on average answering statements about their child expressing internalising problems with “not true” (see Item mean internalising problems in Table 5.2). The pattern of mothers reporting almost no internalising problems and similar levels of internalising problems for female and male offspring did not change in the samples with 7-year-old children (Table 5.2).

Table 5.2 Offspring characteristics

Cohort (N)	M age in months (SD)*	M gestational age in weeks (SD)	N (~%) Female	Item mean (SD) of internalising subscale**	Item mean (SD) of internalising subscale ** – Female sex	Item mean (SD) of internalising subscale ** – Male sex
Internalising problems age 3 ~ Cord blood DNAm						
ALSPAC (N = 739)	47.59 (0.85)	39.55 (1.49)	377 (51%)	0.26 (0.27)	0.28 (0.04)	0.25 (0.28)
Generation R (N = 793)	36.36 (0.84)	40.26 (1.40)	400 (50%)	0.12 (0.10)	0.12 (0.11)	0.12 (0.10)
Moba1 (N = 998)	47.59 (0.85)	39.97 (1.55)	499 (50%)	0.24 (0.42)	0.25 (0.22)	0.23 (0.21)
Moba 2 (N = 481)	28.93 (16.38)	39.95 (1.45)	212 (44%)	0.24 (0.22)	0.25 (0.24)	0.23 (0.20)
Mean or total (N = 3,011)	41.67 (1.18)	39.92 (1.50)	1445 (48%)	-	-	-
Internalising problems age 7 ~ Cord blood DNAm						
ALSPAC (N = 709)	81.20 (1.00)	39.54 (1.51)	356 (50%)	0.25 (0.28)	0.27 (0.27)	0.24 (0.29)
Generation R (N = 892)	70.92 (3.6)	40.25 (1.40)	448 (50%)	0.14 (0.14)	0.13 (0.13)	0.14 (0.14)
Mean or total (N = 1,601)	80.29 (1.43)	39.91 (1.49)	804 (50%)	-	-	-
Internalising problems age 7 ~ Peripheral blood DNAm at age 7						
ALSPAC (N = 747)	81.22 (1.06)	-	372 (50%)	0.27 (0.29)	0.27 (0.27)	0.26 (0.30)
Generation R (N = 374)	71.28 (4.32)	-	195 (52%)	0.12 (0.13)	0.12 (0.12)	0.13 (0.14)
Mean or total (N = 1,121)	80.93 (1.28)	-	572 (51%)	-	-	-

*Note. M = mean. In the Total row, means were weighted by the inverse variance for each cohort. * age at DNA methylation assessment. ** As each cohort used different scales, it was not possible to generate a total sample mean for internalising problems. DNAm = DNA methylation. Item means were calculating by dividing the total score by the number of items.*

5.4.1.3 Associations between offspring internalising problems and covariates

Across cohorts and time points there was statistical evidence for maternal anxiety and depressive symptoms during pregnancy being positively associated with offspring internalising problems during childhood, yet the magnitude of associations was rather small (Table 5.3). As expected, there was evidence for internalising problems assessed at the age of 3 being positively associated with internalising problems assessed at the age of 7 but still only of moderate size, indicating that internalising problems change from age 3 to 7 (Table 5.3). Other than these associations, and a small association between younger maternal age at delivery and increased offspring internalising problems at the age of 3 in the MoBa1 cohort, there was little evidence of association between offspring internalising problems and any of the other covariates (Table 5.3). Offspring of mothers who smoked throughout or late during pregnancy had similar mean internalising problems at age 3 and 7, as offspring whose mothers quit early or never smoked during pregnancy (Table 5.4).

Table 5.3 Correlation table of internalising problems and covariates

	Correlation with int. problems age 3				**Correlation with int. problems age 7**	
Covariates	*ALSPAC*	*Generation R*	*MoBa1*	*MoBa2*	*ALSPAC*	*Generation R*
Maternal anxiety/depression during pregnancy	0.13**	0.23**	0.19**	0.19**	0.15**	0.20**
Maternal education	-0.03	-0.06	-0.01	-0.02	0.05	-0.03
Maternal age	-0.05	-0.02	-0.07*	-0.08	-0.01	-0.07*
Gestational age	-0.02	0.01	-0.06	-0.05	-0.03	-0.05
Internalising problems age 3	-	-	-	-	0.39**	0.56**

*Note. int problems = internalising problems. * correlation statistically different from zero (P-value < 0.05). ** correlation statistically different from zero (P-value < 0.001).*

Table 5.4 Comparison of internalising problems between offspring exposed/ not exposed to smoking during pregnancy

	M int. problems age 3 (SD)				**M int.problems age 7 (SD)**	
Maternal smoking during pregnancy	*ALSPAC*	*Generation R*	*MoBa1*	*MoBa2*	*ALSPAC*	*Generation R*
No smoking or quitting early in pregnancy	0.25 (0.27)	0.12 (0.11)	0.23 (0.22)	0.23 (0.22)	0.25 (0.28)	0.14 (0.14)
Smoking late/throughout pregnancy	0.26 (0.29)	0.12 (0.10)	0.25 (0.19)	0.29 (0.23)	0.26 (0.29)	0.14 (0.12)

Note. M = mean int problems = internalising problems.

5.4.2 Probe-level meta-analysis

5.4.2.1 Quality control checks for cohort results

The correlation matrices of the individual cohort results can be found in Appendix U, Figures U1 to U8. In each cohort, the effect estimates of each of the models correlated moderately to highly and the pattern of correlations was consistent across cohorts, assessment time-points, and tissue of DNA methylation (cord vs. peripheral blood), indicating that each model was assessing the same effect. Lowest correlations were observed between the effect estimates of the sex-stratified models, which may be explained by the small sample size in the stratified analyses, residual confounding (Yousefi et al., 2015), or sex-differences between the associations of DNA methylation and internalising problems. As most of the P-values in the QQ-plots follow the null line closely and stay within the 95% Confidence Intervals, there is little evidence for systematic confounding of the association between DNA methylation at individual CpG sites and offspring internalising problems (Appendix V, Figures V1 to V8). The precision plots show that ALSPAC has the highest precision in the prenatal analyses, despite MoBa1 and Generation R having slightly larger sample sizes (Appendix W, Figures W1 and W2. In the peripheral blood analysis, ALSPAC showed higher precision than Generation R, as would be expected because of the larger sample size (Appendix W, Figure W3).

5.4.2.2 Quality control checks for meta-analysed results

Correlation matrices of effect estimates of the models of the meta-analysed results were in line with the correlations observed in the individual cohorts (Appendix X, X1 to X3). Overall, estimates correlated highly, except for the sex stratified model coefficients, which only correlated moderately with the other model estimates, and did not show any association with each other. This pattern was consistently observed within cohorts, across tissues, as well as in the meta-analysed results, and could indicate potential sex-differences in the association between DNA methylation and internalising problems. Alternatively, this may also be explained by residual confounding of sex (Yousefi et al., 2015) and/or the small sample size in the sex-stratified analyses. Spurious results in sex-stratified EWAS models have previously been reported in a study where offspring sex was less likely to be a confounder/important source of variation in the trait of interest (paternal BMI)

(Sharp et al., 2021). Visual inspection of QQ-plots indicates that the P-values are mostly normally distributed (Appendix Y, Figures Y1 to Y3). The plots of the leave-one-out analysis can be found in Appendix Z, Figures Z1 to Z3).

5.4.2.3 Association between DNA methylation and offspring internalising problems at the age of 3 and 7

An overview of the results of the probe-level meta-analysis of each model can be found in Table 5.5. The cord blood meta-analysis using internalising problems at the age of 3 did not show any evidence for a statistically significant association between DNA methylation in cord blood and internalising problems. In the cord blood and childhood peripheral blood analyses using internalising problems at the age of 7 as the outcome, there was weak evidence for DNA methylation being associated with internalising problems in female sex offspring only (Table 5.5 and Table 5.6). The cord blood meta-analysis showed evidence for one CpG site being significantly associated with internalising problems in female sex offspring at the age of 7 (Chr6: cg266686320; estimate = -8.76, SE = 1.54, FDR adjusted P-value = 0.005), which did survive the leave-one-out analysis (Appendix G, Figure G1).

Also, the childhood peripheral blood analysis did find evidence for DNA methylation at two CpG sites being significantly associated with internalising problems in female sex offspring at the age of 7 (Chr6: cg07283896; estimate = 6.10, SE = 0.97, FDR adjusted P-value = 0.0001 and Chr7: cg08884410; estimate = 8.59, SE = 1.62, FDR adjusted P-value = 0.02; Table 1.5 and Table 1.6). However, of those CpG sites only cg08884410 survived the leave-one-out analysis (Appendix Z, Figures Z2 and Z3). The estimates are reflecting the change in internalising z-score per an increase from 0% to a 100% DNA methylation at a given CpG site. As DNA methylation does not change from 0% to 100% methylated or vice versa (see Chapter 2) (Mulder et al., 2020), a more reasonable interpretation would be to divide the estimate by 10, so that the estimate is representing a change in internalising z-score per 10% increase in percentage DNA methylation points at a given CpG site. For the cord blood meta-analysis, a 10-point increase in percentage DNA methylation at cg266686320 would therefore be associated with a 0.9 decrease in offspring's internalising z-scores at the age of 7 (Table 5.6). In the cross-sectional analysis, estimates were of similar magnitude; with a 10-point increase in percentage DNA methylation at

cg08884410 being associated with a 0.9 increase in female sex offspring's internalising problems at the age of 7.

Table 5.5 A summary of results of each EWAS model from the probe-level analysis

Model	N CpG sites with FDR-corrected P-value <0.05	N CpG sites (%) surviving leave-one-out analysis	Meta-analysis sample size	Genomic inflation factor (λ)**
Internalising problems age 3 ~ Cord blood DNA methylation (N = 3,011)				
All offspring (minimally adjusted)*	0	n.a.	3011	1.01
All offspring (adjusted for covariates)	0	n.a.	3011	1.00
Female sex offspring (adjusted for covariates)	0	n.a.	1452	1.02
Male sex offspring (adjusted for covariates)	0	n.a.	1559	0.99
All offspring (adjusted for covariates and maternal anx/dep)	0	n.a.	2987	1.00
Internalising problems age 7 ~ Cord blood DNA methylation (N = 1,601)				
All offspring (minimally adjusted)*	0	n.a.	1601	1.02
All offspring (adjusted for covariates)	0	n.a.	1601	1.03
Female sex offspring (adjusted for covariates)	1	1 (100%)	804	1.02
Male sex offspring (adjusted for covariates)	0	n.a.	797	1.06
All offspring (adjusted for covariates and maternal anx/dep)	0	n.a.	1601	1.03
Internalising problems age 7 ~ Childhood DNA methylation age 7 (N = 1,121)				
All offspring (minimally adjusted)*	0	n.a.	1601	0.99
All offspring (adjusted for covariates)	0	n.a.	1136	0.98
Female sex offspring (adjusted for covariates)	2	1 (50%)	572	1.02
Male sex offspring (adjusted for covariates)	0	n.a.	564	0.94
All offspring (adjusted for covariates and maternal anx/dep)		n.a.	1136	0.98

*Note. * only adjusted for estimated cell counts and 20 surrogate variables. Covariates: maternal age, maternal smoking, maternal education, offspring age, 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.*

Table 5.6 CpG sites associated with offspring DNA methylation and internalising problems at the age of 7 with FDR-adjusted P-value < 0.05

Predictor	CpG (gene)	Estimate (SE)	P-value	BF adjusted P- value
Internalising problems age 7 ~ Cord blood DNA methylation (Female sex only)	Chr6: cg26668632 (IFNGR1)	-8.76 (1.54)	1.42×10^{-08}	0.005
Internalising problems age 7 ~ Childhood DNA methylation age 7 (Female sex only)	Chr6: cg07283896*	6.10 (0.97)	3.20×10^{-10}	0.0001
	Chr7: cg08884410	8.59 (1.62)	1.11×10^{-07}	0.02

*Note. "Estimate" can be interpreted as difference in internalising z-score per one unit increase in percentage DNA methylation points, after adjustment for all covariates. BF = Bonferroni. * did not survive the leave-one-cohort-out analysis.*

5.4.3 Differentially methylated regions meta-analysis

5.4.3.1 Cord blood meta-analyses

5.4.3.1.1 Internalising problems at the age of 3

The DMRff meta-analysis found evidence for 2 to 3 differentially methylated regions (DMRs) in cord blood, with a length of 3 to 4 CpG sites to be associated with internalising problems at the age of 3 (Table 5.7). The covariate adjusted model showed evidence for two DMRs containing CpG sites annotated to the *STK32C* (Chr10: 134045514-134045913; estimate = 5.08, SE = 0.88; P-value = 9.230×10^{-09}) and *MIR886* gene (Chr5: 135415693-135416613; estimate = -4.80, SE = 0.87; P-value = 3.948×10^{-08}), of which only the region annotated to *STK32C* gene survived adjustment for maternal anxiety and depressive symptoms during pregnancy (Chr10: 134045514-134045913; estimate = 5.37, SE = 0.88, P-value = 1.165×10^{-09}). No regions were identified in the covariate models stratified by male sex, but three regions were identified in the female stratified models, which is in line with the probe-level results and indicates more evidence for DNA methylation being associated with internalising problems in female sex offspring. One of these regions overlapped with the region found in the covariate-adjusted model, which was annotated to the MIR886 gene (Chr5: 135415693-135416613; estimate = -8.54, SE = 1.28, P-value = 2.870×10^{-11}).

Table 5.7 Results from the meta-analysis of differentially methylated regions for the cord blood meta-analysis: Internalising problems age 3

DMR	N CpG sites	Estimate (SE)	Z	P-value	BF adjusted P-value	Gene
Covariate adjusted model						
Chr10: 134045514-134045913	4	5.08 (0.88)	5.74	9.230×10^{-09}	0.003	*STK32C*
Chr5: 135415693-135416613	15	-4.80 (0.87)	-5.49	3.948×10^{-08}	0.014	*MIR886*
Covariate adjusted model – Stratified by female sex						
Chr5: 135415693-135416613	15	-8.54 (1.28)	-6.65	2.870×10^{-11}	1.049×10^{-05}	*MIR886*
Chr7: 95025611-95026095	13	39.50 (7.18)	5.50	3.819×10^{-08}	0.014	*PON3*
Chr6: 32120783-32121055	10	-17.66 (3.24)	-5.46	4.868×10^{-08}	0.018	*PPT2; PRRT1*
Covariate adjusted model – Stratified by male sex						
	0					
Covariate adjusted model –Maternal anx/dep adjusted						
Chr10: 134045514-134045913	4	5.37 (0.88)	6.08	1.165×10^{-09}	0.0004	*STK32C*

Note. DMR = differentially methylated region. BF = Bonferroni. Covariate adjusted models were adjusted for maternal age, maternal smoking, maternal education, offspring age, estimated cell counts and 20 surrogate variables.

5.4.3.1.2 Internalising problems at the age of 7

The meta-DMR analysis in cord blood showed evidence DNA methylation at 3 to 4 regions being associated with internalising problems at the age of 7 (Table 5.8), yet none of these regions were the same as the regions found at age 3 (Table 5.7). In the covariate adjusted model one region annotated to gene *PFKFB2* was found (Chr1: 207226769-207226830; estimate = 196.12, SE = 34.04, P-value = 8.377 x 10^{-09}) and this region also survived additional adjustment for maternal anxiety and depression symptoms during pregnancy (Chr1: 207226769-207226830; estimate = 196.05, SE = 33.67, P-value = 5.805 x 10^{-09}). Two regions were detected in each of the sex-stratified models, which were not overlapping (Table 5.8). Also, the regions associated with internalising problems of female sex offspring at the age of 3 were not replicated at age 7. The two regions found in the female-sex stratified model were annotated to *MIR548F5; NBEA; MAB21L1* (Chr13: 36050844-36050889; estimate = -79.94, SE = 14.73, P-value = 5.775 x 10^{-08}) and *SLC1A2* (Chr11: 35441311-35441777; estimate = -456.04, SE = 84.32, P-value = 6.359 x 10^{-08}). Only one of the two regions of the male-sex stratified model was annotated to a gene, namely *ADSSL1* (Chr14: 105190337-105190477, estimate = 648.61, SE = 108.59, P-value = 2.330 x 10^{-09}). The non-annotated region was found on chromosome 6, spanning three CpG sites (Chr6: 75918463-75918733, estimate = 205.00, SE = 38.09, P-value = 7.370 x 10^{-08}).

Table 5.8 Results from the meta-analysis of differentially methylated regions for the cord blood meta-analysis: Internalising problems age 7

DMR	N CpG sites	Estimate (SE)	Z	P-value	BF adjusted P-value	Gene
Covariate adjusted model						
Chr1: 207226769-207226830	3	196.12 (34.04)	5.76	8.377×10^{-09}	0.003	*PFKFB2*
Covariate adjusted model – Stratified by female sex						
Chr13: 36050844-36050889	3	-79.94 (14.73)	-5.43	5.775×10^{-08}	0.022	*MIR548F; NBEA; MAB21L1*
Chr11: 35441311-35441777	4	-456.04 (84.32)	-5.41	6.359×10^{-08}	0.024	*SLC1A2*
Covariate adjusted model – Stratified by male sex						
Chr14: 105190337-105190477	3	648.61 (108.59)	5.97	2.330×10^{-09}	0.001	*ADSSL1*
Chr6: 75918463-75918733	3	205.00 (38.09)	5.38	7.370×10^{-08}	0.023	-
Covariate adjusted model –Maternal anx/dep adjusted						
Chr1: 207226769-207226830	3	196.05 (33.67)	5.82	5.805×10^{-09}	0.002	*PFKFB2*

*Note. * only adjusted for estimated cell counts and 20 surrogate variables. Covariates: maternal age, maternal smoking, maternal education, offspring age, 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. DMR = differentially methylated region. BF = Bonferroni.*

5.4.3.2 Childhood peripheral blood analyses

5.4.3.2.1 Internalising problems at the age of 7

The DMR meta-analysis found evidence for 2-5 differentially methylated regions, spanning 2-10 CpG sites, in offspring's peripheral blood at the age of 7 that were associated with internalising problems at the age of 7 (Table 5.9). In the covariate adjusted model, two regions were found that were mapped to the *C10orf26* (Chr10: 104535961-104536121; estimate = -55.48, SE = 8.83, P-value = 3.302 x 10^{-10}) and *FAM125A* (Chr19:17531148-17531370, estimate = 154.00, SE = 28.79, P-value = 8.832 x 10^{-08}) gene. Of these two regions, none survived adjustment for maternal anxiety and depression during pregnancy (Table 5.8). The maternal anxiety and depression adjusted model found one DMR, with 3 CpG sites, mapped to the *FAM848A* gene to be associated with offspring internalising problems (Chr2:14772562-14772568; estimate = 5.37; SE = 14.21; P-value = 8.049 x 10^{-08}). The sex-stratified models showed evidence for 5 DMRs in the female sex-, and only 1 DMR in the male-sex stratified models, to be associated with internalising problems. Four regions of the female stratified model were annotated to the *MIR2110; C10orf118* gene (Chr10: 115934483-115934488, estimate = 535.78, SE = 92.54, P-value = 7.046^{-09}), *HMGB1* (Chr13: 31039916-31040215, estimate = 219.62, SE = 38.51, P-value = 1.178 x 10^{-08}), FBXO21 (Chr12: 117627133-117627450, estimate = 85.11, SE = 15.81, P-value = 7.299 x 10^{-08}) and *KCNJ9* (Chr1: 160054219-160054321; estimate = 52.88, SE = 9.99, P-value = 1.188 x 10^{-07}) gene. The largest region associated with female offspring internalising problems at the age of 7 was on chromosome 6, spanning 10 CpG sites but has no gene annotated (Table 5.9). The male stratified model showed 1 DMR, spanning eight CpG sites, that was annotated to the *HOXA4* gene (Chr7: 27170412-27170994; estimate = -18.33, SE = 2.65; P-value = 4.814 x 10^{-12}).

Table 5.9 Results from the meta-analysis of differentially methylated regions for the cross-sectional analysis: Internalising problems age 7

DMR	N CpG sites	Estimate (SE)	Z	P-value	BF adjusted P-value	Gene
Covariate adjusted model						
Chr10:104535961-104536121	5	-55.48 (8.83)	-6.28	3.302 x10^{-10}	1.233^{-04}	C10orf26
Chr19:17531148-17531370	3	154.00 (28.79)	5.35	8.832 x10^{-08}	0.033	FAM125A
Covariate adjusted model – Stratified by female sex						
Chr10:115934483-115934488	2	535.78 (92.54)	5.79	7.046 x10^{-09}	0.022	MIR2110; C10orf118
Chr13:31039916-31040215	7	219.62 (38.51)	5.70	1.178 x10^{-08}	0.003	HMGB1
Chr6:31148370-31148552	10	14.56 (2.57)	5.67	1.419 x10^{-08}	0.004	-
Chr12:117627133-117627450	4	85.11 (15.81)	5.38	7.299 x10^{-08}	0.005	FBXO21
Chr1:160054219-160054321	3	52.88 (9.99)	5.30	1.188 x10^{-07}	0.028	KCNJ9
Covariate adjusted model – Stratified by male sex						
Chr7:27170412-27170994	8	-18.33 (2.65)	-6.91	4.814 x10^{-12}	1.789 x10^{-06}	HOXA4
Covariate adjusted model –Maternal anx/dep adjusted						
Chr2:14772562-14772568	3	76.23 (14.21)	5.37	8.049 x10^{-08}	0.030	FAM84A

*Note. * only adjusted for estimated cell counts and 20 surrogate variables. Covariates: maternal age, maternal smoking, maternal education, offspring age, 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. DMR = differentially methylated region. BF = Bonferroni.*

5.4.4 Candidate gene look-up

In addition to the hypothesis-free epigenome-wide approach, a targeted gene approach was performed for the four candidate genes described in the introduction. An overview of these results is presented in the Manhattan plots in Figure 5.3 to Figure 5.5). The CpG sites that reached nominal significance were checked for consistency in their direction of effect and P-value in the cord blood and peripheral blood analysis at the age of 7 and results are presented in Appendix BB, Table BB1).

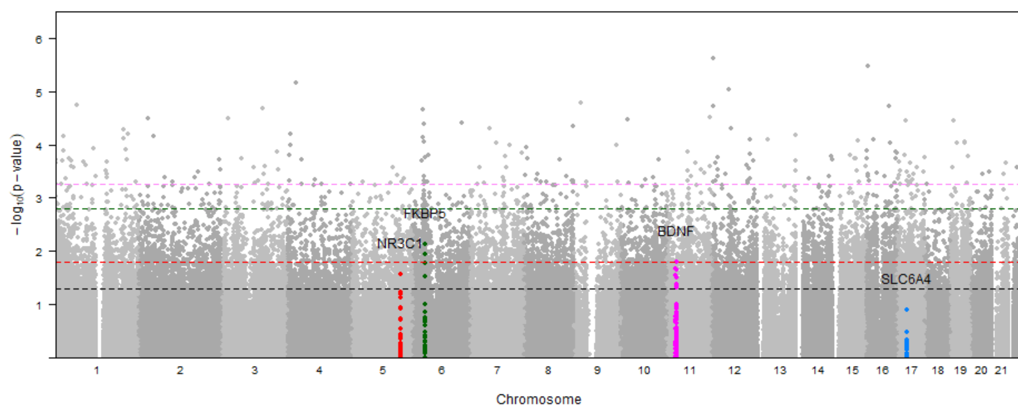


Figure 5.3 Manhattan plot of the candidate gene look-up of the internalising problems age 3 ~ cord-blood results. *The black line represents the nominal significance level. The coloured lines represent the corresponding Bonferroni threshold (matched by colour to the corresponding gene).*

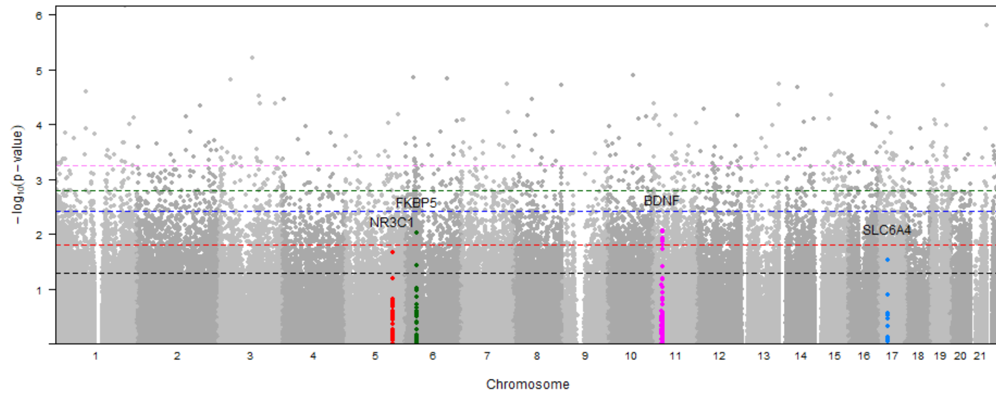


Figure 5.4 Manhattan plot of the candidate gene look-up of the internalising problems age 7 ~ cord-blood results. *The black line represents the nominal significance level. The coloured lines represent the corresponding Bonferroni threshold (matched by colour to the corresponding gene).*

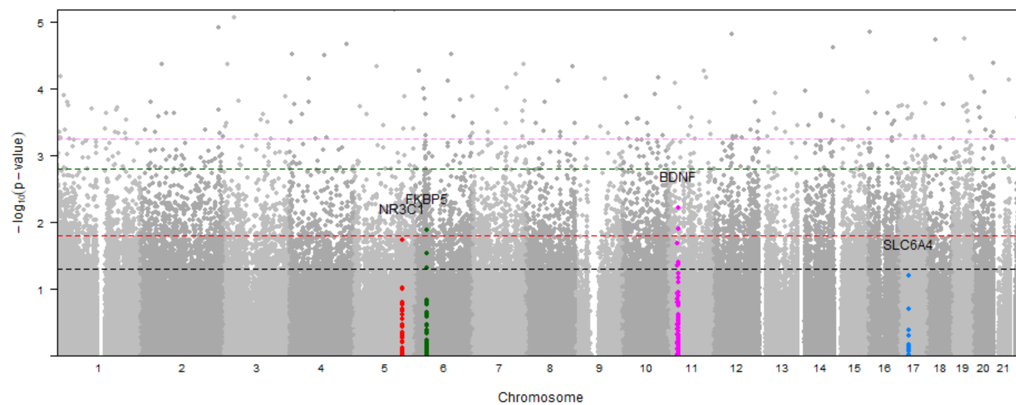


Figure 5.5 Manhattan plot of the candidate gene look-up of the internalising problems age 7 ~ childhood peripheral blood results. *The black line represents the nominal significance level. The coloured lines represent the corresponding Bonferroni threshold (matched by colour to the corresponding gene).*

There were 32 CpG sites annotated to the candidate gene *NR3C1*, with one CpG site reaching nominal significance in the cord blood analysis of internalising problems at the age of 3 (Chr5: cg19457823, estimate = -1.17, SE = 0.53, P-value = 0.03) but not falling below the Bonferroni adjusted P-value threshold of 0.002 (0.05/32 tests; red line, Figure 5.3), and not showing a consistent direction of effect in the cord blood analysis using internalising problems at the age of 7 (Appendix BB, Table BB1). Overall, this indicates little evidence that DNA methylation at *NR3C1* is contributing towards the development of internalising problems in children in this sample. Likewise, there was no evidence of DNA methylation at *SLC6A4* gene influencing the development of childhood internalising problems. None of the 13 *SLC6A4* annotated CpG sites even reached nominal significance (Figure 5.3) in the cord blood meta-analysis at the age of 3. Out of 91 CpG sites annotated to *BDNF*, only 9 reached nominal significance (black line, Figure 5.3) and none survived adjustment for multiple testing using the Bonferroni correction (Bonferroni P-value threshold: 0.05/91 tests = 0.0005; purple line, Figure 5.3). Out of the 9 CpG sites with nominal significance, only 1 CpG site showed a consistent direction of effect in all three analyses, as well as a nominal significant P-value in the cord blood meta-analysis using internalising problems at the age of 7 (Chr11: cg10022526, estimate = -4.24, SE = 1.76, P-value = 0.016; Appendix I, Table I1). However, the effect appeared to be very small in the childhood peripheral blood analysis using internalising problems at the age of 7 (estimate = -0.74, SE = 2.58, P-value = 0.774; Appendix BB, Table BB1). This provides weak evidence that DNA methylation at cg10022526 might be associated with prenatal exposures and internalising problems in childhood. Strongest evidence was found for DNA methylation at the *FKBP5* gene being associated with internalising problems, with 5 out of 32 annotated CpG sites in the prenatal age 3 analysis reaching nominal significance (black line, Figure 5.3), of which 4 showed a persistent direction of effect (Appendix BB, Table BB1) but no P-value below the Bonferroni adjusted thresholds (Figure 5.3 to Figure 5.5; green line). In summary, there was no evidence for DNA methylation at 4 candidate genes, which have previously been associated with offspring's stress response, being associated with childhood internalising problems in this sample.

5.4.5 Meta-analysis of cell proportions

In a meta-analysis of associations between internalising disorders and estimated cell proportions, neither the cord blood cell proportions nor the peripheral blood cell proportions were associated with offspring internalising problems (Appendix AA, AA1 to AA18).

5.5 Discussion

5.5.1 Summary and interpretation of findings

Overall, the results of this study provide weak evidence for an association between DNA methylation and internalising problems in childhood. First, there were only few associations between DNA methylation at individual probes and regions, and internalising problems at the age of 3 and 7 years. Second, there was no overlap between CpG sites or regions across different assessment time-points of DNA methylation and internalising problems.

The only pattern that could be observed across assessment time points was more evidence for associations with internalising problems in female than male sex children. The probe-level analysis showed weak evidence for DNA methylations at one CpG site in cord blood and two CpG sites in peripheral blood (of which only one survived the leave-one-out analysis) to be associated with internalising problems at the age of 7, in female sex children. The pattern of the probe-level analysis was confirmed by the DMR analyses, in which the highest number of internalising problem-associated DMRs was observed for female sex children in the cross-sectional analysis. Recent research indicates little systematic sex difference (~5%) in the development of DNA methylation at autosomes from birth to age 6 (Mulder et al., 2020), with female offspring tending to show higher levels of DNA methylation than male offspring. This could explain some, but not all of the associations that were observed in the probe-level and regional analysis. In the literature, internalising problems are reported to be more prevalent in female sex children than male sex (Gutman & Codiroli McMaster, 2020). Yet, it is unclear whether this sex difference in internalising problems is due to different biological pathways, socialisation, or both (Gutman & Codiroli McMaster, 2020).

There are several potential explanations for the missing overlap between the probes and regions that were found to be associated with internalising problems at

age 3 and 7: (1) the effects are not persistent over development and/or of small magnitude, or (2) represent false positives/negatives due to the difference in sample size (and therewith power) between the age 3 and age 7 analysis. Alternatively, (3) it could also indicate that different developmental stages are subject to different confounding structures. Last, the use of different tissue types may have contributed to the lack of overlap in results since the comparison of DNA methylation in cord blood with peripheral blood in childhood might be problematic because of differences in cell compositions and gene regulations between these tissues (Martino et al., 2011; Mulder et al., 2020). However, this cannot complete account for the missing congruence, as results of this study did not show more alignment between DNA methylation in the two cord blood analyses than in the cord blood and peripheral blood analyses.

The effect sizes of the probe-level EWAS meta-analysis appeared to be of similar magnitude as effects found in similar studies of childhood mental health problems, such as ADHD (Neumann et al., 2020) and aggressive behaviour (Dongen et al., 2021), as well as more distantly related phenotypes, such as BMI. In the EWAS meta-analysis of ADHD (N = 2,477) the effect sizes of the top hits showed a 10% increase in DNA methylation points at birth to be associated with changes in children's ADHD z-scores of 0.2 to 0.4 (Neumann et al., 2020). In the EWAS meta-analysis of childhood aggressive behaviour (N = 2,425) the top site showed a 10% increase in DNA methylation points at birth to be associated with a 0.4 z-score increase in childhood aggressive behaviour (Dongen et al., 2021). The top CpG site at birth of the EWAS meta-analysis of childhood BMI showed a sex- and age-adjusted standard deviation score of 0.27 (SE = 0.05; N = 3,295) between the ages 2 to 5 years, and of 0.96 (SE = 0.17; N = 4,133) at 5 to 10 years, in association with 10% increase in methylation points (Vehmeijer et al., 2020). In comparison, in this study a 10% increase in DNA methylation points at the top site at birth in female sex children was associated with a decrease in internalising problems z-score of 0.9. The effect sizes of the top CpG sites of the non-stratified models ($P < 1x10^{-5}$, Appendic CC, Table CC1) at birth showed a range of internalising problems z-scores of 0.1 to 1.1, and in the cross-sectional analysis of 0.2 to 0.5, in association with a change in 10% methylation points (Appendix CC, Table CC1).

5.5.2 Biological pathways

According to the database look-up, some of the differentially methylated regions that were found to be associated with childhood internalising problems were annotated to genes linked to related traits such as: Brain development (*PPT1*, *MAB21L1*) and neurological functioning (*SCL1A2*, *NBEA*, *PPT2*) and previously associated with anxiety-related disorders (*NBEA, C10orf26/WBP1L*). Additionally, two of the genes annotated to the DMRs have previously been associated with autism spectrum disorder (*PRRT1*, *NBEA*) and schizophrenia (*PRRT1*, *NBEA*, *C10orf26/WBP1L*). However, most of the DMR annotated genes have not been linked to traits related to internalising problems (*MIR886*, *PON3*, *PFKFB2*, *ADSSL1*, *HMGB1*, *FBXO21*, *KCNJ9*, *FAM125A*). Interestingly, the *STK32C* gene that was annotated to a DMR of the cord-blood age 3 analysis has been reported to show increased buccal cell DNA methylation in a small EWAS of 18 monozygotic twins discordant for adolescent depression (Dempster et al., 2014). Also, this was replicated in post-mortem brain tissue of veterans discordant for depression (Dempster et al., 2014). Furthermore, in a different study, a closely related gene region near the promoter of the *STK32B* gene was differentially methylated in adolescents at risk to develop generalised anxiety disorder (GAD; N = 221) (Ciuculete et al., 2018). However, the effect was very small, with adolescents high at risk for GAD only showing 1% increased methylation at a CpG site annotated to *STK32B,* compared to adolescents at low risk for GAD.

However, as the *STK32C* region did not come up in the cord blood and peripheral blood age 7 analyses, it is not clear whether this a true positive finding or due to chance. Generally, the little congruence between results of the different models and time-points should alert to refrain from any biological interpretations before future research has replicated these findings.

Also, the candidate-wide analyses did not support evidence for a contribution of DNA methylation differences in cord blood at *NR3C1*, *SLC6A4*, *BDNF* or *FKBP5* to the risk for internalising problems in childhood. Only a few of the CpG sites annotated to these genes showed nominal significant associations with internalising problems and none survived the Bonferroni correction (Figure 5.3 to Figure 5.5 and Appendix BB, Table BB1). However, most of the previous candidate gene studies were applied in samples of children with exposure to

adverse experiences, which was not factored into the analyses of this study. For instance, Weder and colleagues (Weder et al., 2014) did not find evidence for differential DNA methylation being associated with depression in 190 children (aged 5-14 years) in a candidate gene analysis of the same genes used in this study. However, they did find differences between maltreated and non-maltreated children at few probes that survived Bonferroni correction (N probes *NR3C1* = 1; N probes *BDNF* = 9*;* N probes *FKBP5* = 4), indicating that DNA methylation at these genes may be uniquely responsive to maltreatment (Weder et al., 2014). The demographics of this sample indicate a rather low risk for children to be exposed to adverse experiences (high maternal education, few mothers smoking during pregnancy, few mothers with anxiety/depression during pregnancy) and low levels of offspring internalising problems across the cohorts. Thus, results might be different in higher risk populations. Further, it has been noted that the 450k array, which has been used to assess DNA methylation in this study, does not assess methylation at the same sites as most of the candidate studies have (Weder et al., 2014), making direct replication of DNA methylation at these sites impossible in this study.

5.5.3 Strengths

The present investigation of the association between DNA methylation and the development of internalising problems in very young children complements the current evidence base on the involvement of DNA methylation in the development of anxiety and depressive disorders, which predominately comprises of adult populations and symptoms measured later in life (Barker, Walton, & Cecil, 2018; Li et al., 2019). Moreover, the investigation of associations with two different assessment time points of DNA methylation (cord- and peripheral-blood), enabled to evaluate the importance of prenatal and postnatal risk factors, and the investigation of two different assessment time-points of internalising problems (age 3 and age 7) allowed evaluating the persistence of these effects across development. Another strength is in the application of a genome-wide probe-level and regional analysis, which allowed exploring the contribution of DNA methylation in a hypothesis-free way, which might shed light on novel biological pathways to anxiety and depression. The importance of hypothesis-free investigations of DNA methylation instead of candidate-gene analyses was

highlighted by the candidate gene look-up of this study. Despite higher statistical power due to less stringent multiple testing penalties (Jones et al., 2018), the candidate gene look-up did not show evidence for associations between DNA methylation at these genes and internalising problems. Candidate-gene findings commonly fail to replicate in genome-wide analyses and the usefulness of candidate gene approaches in the era of epi(genome)-wide association studies has been questioned (Barker, Walton, & Cecil, 2018; Duncan et al., 2019; Shabalin & Aberg, 2015). Last, the decent sample size and prospective design of the studies used in the meta-analysis allowed getting an indication for the temporal sequence of events and reduced recall bias.

5.5.4 Limitations

The results of this study should be interpreted by considering the following limitations. First, this study investigated the influence of DNA methylation on internalising problems in very young children. One of the main aims of this study was to understand the influence of environmental exposures during pregnancy on offspring internalising problems by using cord blood DNA methylation as a proxy for the quality of the intrauterine environment. As DNA methylation is highly susceptible to environmental exposures, the internalising problems assessment time point was chosen based on minimizing the temporal gap between exposure and outcome, to reduce the number of confounding factors. However, this came at the costs of the accuracy and clinical significance of assessing internalising problems. Previous research has refrained from including children younger than four years of age in their studies because the SDQ and CBCL questionnaires have not been developed to evaluate behaviour problems in such young children (Goodman & Scott, 1999), which is confirmed by the internalising scales lacking internal consistency in children under the age of 4 (Maurice-Stam et al., 2018). However, in this study no large difference in scores of internalising problems was observed at the age of 3 compared to the age of 7. Second, statistical power was compromised by the high zero inflation of internalising problems at both assessment time-points and thus, small effects might have been missed. Small effect sizes are common in EWAS studies of mental health problems in children and even larger sample sizes may be required to enable detection of numerous small signals (Neumann et al, 2020). Third, as suggested by their demographic

profile, children were probably at a low risk for being exposed to adverse experiences and thus results are unlikely to generalize to less advantaged populations. Selection bias is common problem in longitudinal cohort studies (see Chapter 2) (Boyd et al., 2013; Taylor, Jones, et al., 2018). Specifically, families with a less advantaged background and with more parental and offspring mental health problems have been found to be more likely to drop-out from longitudinal cohort studies (Wolke et al., 2009), which is partly confirmed by children in this study expressing very low amounts of internalising problems. This selection bias is likely to be even stronger for the DNA methylation subsamples included in this study, which have been selected based on completeness of other variables in the cohorts (see Chapter 2). Fourth, it is possible that measurement error was introduced by using an internalising problems total score, which combines anxiety and depressive symptoms in children. Previous research indicates that anxiety and depression symptoms during childhood have different underlying genetic and environmental structures and do not become strongly associated until adolescence (Waszczuk et al., 2014). Thus, measures differentiating between anxiety and depression symptoms during childhood might have been more adequate for this study than a combined measure. Last, this study relied on DNA methylation assessed in blood, yet the tissue of most interest for mental health phenotypes is DNA methylation in the brain. Even though there is some evidence that, at least part of blood DNA methylation can function as a proxy for brain DNA methylation (Davies et al., 2012; Kaminsky et al., 2009; Walton et al., 2016), it is not clear whether the DNA methylation changes associated with childhood internalising problems observed in this study proxy for DNA methylation in brain tissue. Regardless of tissue concordance, results of this study may contribute towards the development of DNA methylation biomarkers for early detection of internalising problems (Walton et al., 2019). However, formal analysis of DNA methylation as a clinical biomarker of internalising problems is warranted, which requires larger samples and different methodological approaches than used in this study (e.g., less stringent adjustment for confounders, training, and validation data sets, etc.; see Chapter 7) (Hüls & Czamara, 2020).

5.5.5 Future studies

Future EWAS with larger sample sizes are needed to clarify the relationship between DNA methylation and development of internalising problems. The higher-powered regional analysis of this study found more evidence of associations between DNA methylation and internalising problems, which warrant replication by independent studies. Specifically, since annotated genes have not been reported by smaller previous EWAS of child/adolescent depression (except for *STK32C* in the prenatal age 3 analysis) (Dempster et al., 2014; Weder et al., 2014). Future studies should aim to reduce noise in anxiety and depression assessments in children by, for instance, including multiple assessment time-points of anxiety and depression across development and ideally drawing on reports from multiple informants (Sallis et al., 2019). Also, future research should try to replicate this study in populations with more diverse demographic characteristics and a higher risk for adverse childhood experiences. Results of this study suggest potential sex differences between associations of DNA methylation and internalising problems, which might deserve further investigation in future studies. Further, future studies might want to investigate an interaction effect between genotype and DNA methylation, as it has been suggested that certain genotypes may be more susceptible to biologically embed environmental effects (Brummelte et al., 2017; Oberlander et al., 2010; Olsson et al., 2010). This is supported by results of a study of Teh and colleagues (Teh et al., 2014), which suggests that around 75% of the most variable DNA methylation regions might be influenced by the interaction of offspring genotype and intrauterine environment. Furthermore, future studies might want to shift focus from predominately investigating the effect of early adverse experiences on DNA methylation, to also investigating resilience factors in children, by exploring the effect of positive early life experiences on DNA methylation and mental health outcomes (Kentner et al., 2019). Lastly, studies might want to investigate associations between DNA methylation and neuroimaging phenotypes (Walton et al., 2020), which might help to better evaluate the biological relevance of DNA methylation findings by posing as an intermediate phenotype to behavioural manifestations of mental health problems.

5.6 Conclusion

This was the first study to investigate the association between differences in DNA methylation and the development of childhood internalising problems. Overall, results of this study indicate that DNA methylation in blood is unlikely to be strongly associated with childhood internalising problems. However, due to the high zero inflation of internalising problems in this sample, it is difficult to evaluate whether results might translate to populations at higher risk for adverse early life experiences. Despite few children expressing internalising problems in this sample, the DMR analyses suggested differences in DNA methylation at regions annotated to genes known to be involved in neurological functioning and development, particularly in female sex offspring. It is likely that these effects would be further exacerbated in populations with more occurrences of internalising problems. As anxiety and depression are highly polygenic disorders (Howard et al., 2019; Jami et al., 2020; Purves et al., 2020), which are likely influenced by the interplay of small effects of many different genes, more EWAS are needed to replicate or invalidate findings of this study, as well as to detect novel differentially methylated regions, which might have been missed by this study.

Chapter 6 – DNA methylation as a mediator of the association between smoking and caffeine consumption during pregnancy and offspring internalising problems

6.1 Chapter overview

In my previous results chapters, I have found no convincing evidence that there is a DNA methylation signature of prenatal caffeine use (Chapter 3) or internalising problems (Chapter 5). Similarly, in the PheWAS chapter (Chapter 4), I showed weak evidence for a causal intrauterine effect of prenatal smoking or caffeine exposure on offspring mental health outcomes, including internalising problems. Taken together, these results provide some evidence against a causal effect of prenatal smoke/caffeine exposure on offspring internalising disorders, including via a mechanism involving DNA methylation. For completeness of assessing the intermediate biomarker-disease association (step 3 of the meet-in-the-middle approach, see Chapter 1), in this chapter I investigate the association between the prenatal smoking- and caffeine-associated DNA methylation changes and offspring internalising problems. To do this, I first conduct an enrichment analysis to explore whether DNA methylation signals from the internalising problems EWAS are enriched for DNA methylation signals from the prenatal caffeine EWAS (Chapter 3) and a large meta-EWAS of prenatal smoking by Joubert and colleagues (N = 6,685) (Joubert et al., 2016). Second, I use two-sample Mendelian Randomisation to assess whether DNA methylation at CpG sites associated with prenatal caffeine (i.e., those identified in Chapter 3) and smoking (Joubert et al., 2016) exposure, and CpG sites observationally associated with internalising problems (i.e., those identified in Chapter 4), are causally related to internalising problems.

6.2 Introduction

As mentioned in Chapter 1, there are theoretical grounds for assuming that DNA methylation mediates the relationship between prenatal exposures and risk for mental health problems (Barker, Walton, & Cecil, 2018; Pingault et al., 2017). Two-step MR has been suggested as a method to investigate DNA methylation as

a mediator for prenatal exposures on offspring outcomes (Davey Smith & Hemani, 2014; Relton & Davey Smith, 2012). However, the application of this method to investigated intrauterine effects involves stringent data requirements. In two-step MR, two different genetic markers are used as separate proxies for the exposure and mediator of interest. In the first step, the prenatal exposure-DNA methylation relationship is estimated through using a genetic instrument for the prenatal exposure. In the second step, a genetic proxy for local DNA methylation associated with the prenatal exposure is used to assess the association between DNA methylation and the outcome of interest. An example of how the effect of intrauterine caffeine exposure on offspring internalising problems could be investigated through two-step MR is illustrated in Figure 6.1. In the first step, the causal effect of prenatal caffeine exposure on DNA methylation is estimated by using a genetic instrument as a proxy for maternal caffeine consumption (e.g., using a PRS as in the PheWAS analysis in Chapter 5 and illustrated in Figure 6.1). In the second step, a cis-SNP is used as a proxy for caffeine-associated DNA methylation changes (Step 2, Figure 6.1).

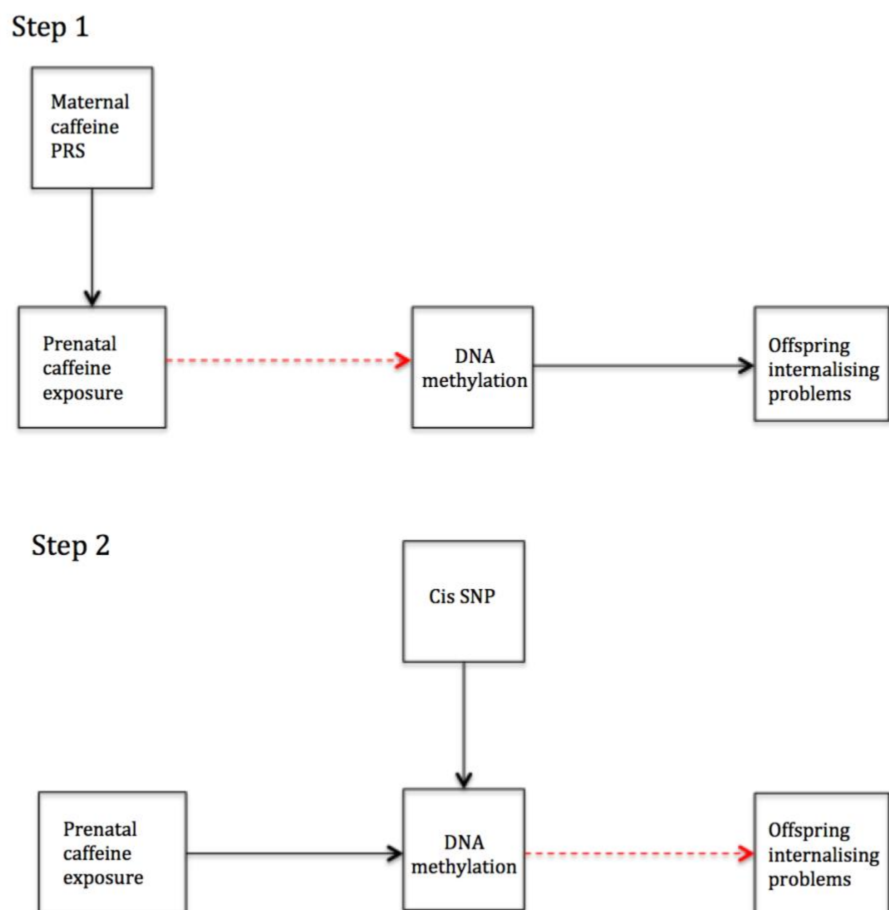


Figure 6.1 Two-step MR to investigate offspring DNA methylation as a mediator of prenatal caffeine exposure on offspring internalising problems. *Step 1: Estimation of the exposure on offspring DNA methylation through using a genetic instrument as a proxy for prenatal caffeine consumption (caffeine PRS). Step 2: Estimation of the association between prenatal caffeine-associated DNA methylation changes and internalising problems in offspring through using a genetic instrument to proxy for caffeine associated CpG sites (cis-SNP) (credit: adapted from Relton & Davey Smith, 2012).*

The first step involves an intergenerational analysis (maternal genetic variants as a proxy for the prenatal exposure on offspring DNA methylation) and thus is subject to the caveats of intergenerational MR, which were described in detail in Chapter 3. Maternal genetic variants capture to 50% offspring genetic variation and therefore, if offspring genotype is not adjusted for, may bias the estimation of the prenatal exposure on offspring outcome (step 1; association observed between the maternal caffeine PRS and offspring DNA methylation may simply represent an association between DNA methylation and offspring PRS). This is problematic for MR as it violates the exclusion restriction assumption of the genetic variant

only being associated with the outcome of interest through the effect of the exposure (Evans et al., 2019). However, as described in Chapter 3, adjustment for offspring genotype may induce collider bias unless the analysis is also adjusted for paternal genotype (see Figure 6.2) (Lawlor et al., 2017). Therefore, in order to derive unbiased estimates, two-step MR requires large sample sizes with data on maternal, offspring, and paternal genotype data, as well as data on offspring DNA methylation and the outcome of interest.

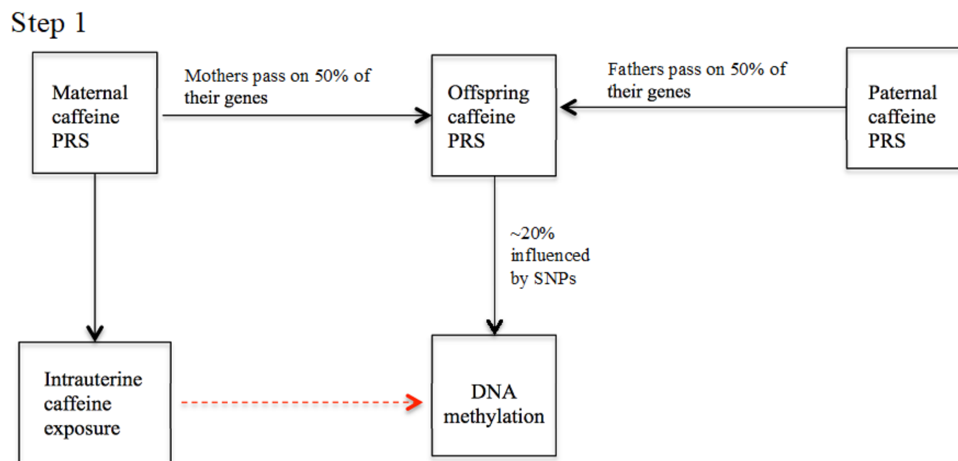


Figure 6.2 Illustration of offspring PRS as a collider of maternal and paternal caffeine PRS. *Red arrow represents to be estimated effect of intrauterine caffeine exposure on offspring DNA methylation. Black arrows represent potential confounding paths.*

If a single sample is able to fulfil these stringent data requirements, an intergenerational two-step, one-sample MR can be conducted. However, such data sources are scarce and thus studies that applied this method tend to be underpowered (Caramaschi et al., 2017). Alternatively, two-step, two-sample MR can be conducted to address power limitations of one-sample MR, by assessing the first and second step in independent samples (Davey Smith & Hemani, 2014). This approach has the great advantage that for the second sample the wealth of GWAS summary data can be exploited to meet the statistical power needed for genetic causal inference techniques (Davey Smith & Hemani, 2014). This is particularly useful for research that investigates molecular pathways, which are commonly restricted to smaller sample sizes (see Chapter 2). Yet, for the intergenerational analysis of step one, GWAS data that has been separated into fetal/maternal components would be required (Evans et al., 2019) – which is currently not available for DNA methylation. For the abovementioned reasons, I was unable to conduct a two-step MR analysis and decided to assess just one-step, in a two-sample MR framework, to investigate the causal effect of prenatal smoking and caffeine associated DNA methylation at birth on offspring internalising problems (in childhood, and anxiety and depression in adulthood). For the one-step, two-sample MR analysis in this chapter, a genetic proxy (cis-SNP) for the exposure-DNA methylation association was retrieved from sample one, which does not require data on the outcome variable of interest (Figure 6.3, top rectangle in red). In a second, independent sample that has genotype data and data on the outcome, the association of the same genetic proxy (cis-SNP) with the outcome is retrieved, without that sample requiring data on DNA methylation (Figure 6.3, bottom rectangle in blue).

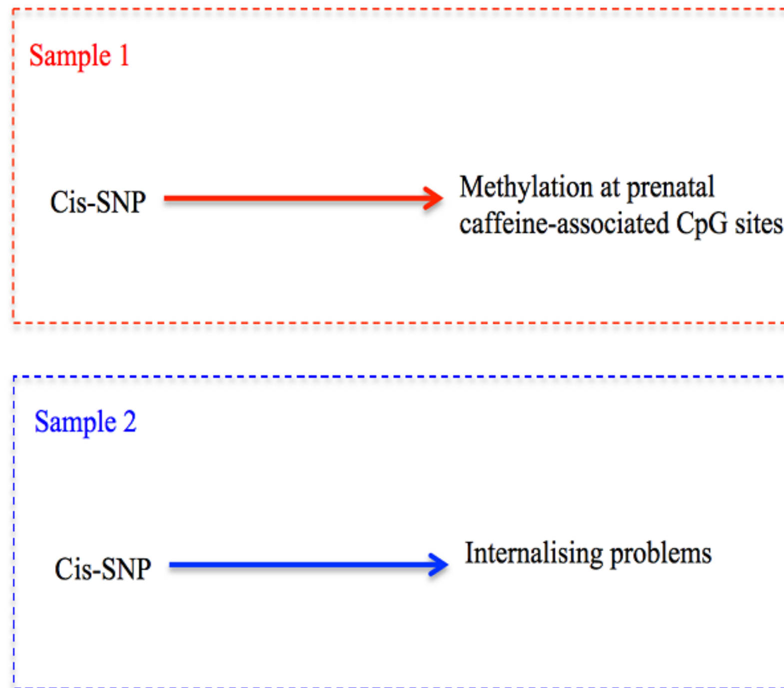


Figure 6.3 Illustration of a two-sample MR analysis to investigate whether DNA methylation mediates the relationship of prenatal caffeine exposure on offspring internalising problems. *Using data of sample 1 (top rectangle in red) the association between a genetic variant on DNA methylation is estimated. Using data of sample 2 (bottom rectangle in blue) the association between a genetic variant that is a proxy for prenatal associated DNA methylation changes and internalising problems is estimated.*

6.2.1 Two-sample MR of prenatal smoking and mental health outcomes

Whilst no study has investigated the causal effect of prenatal caffeine associated DNA methylation on offspring mental health outcomes, several studies have assessed the causal effect of the CpG sites found to be differentially methylated according to maternal smoking during pregnancy (Richardson et al., 2019; Wiklund et al., 2019). Richardson and colleagues (Richardson et al., 2019) conducted a thorough analysis of DNA methylation at the prenatal smoking associated CpG sites from Joubert and colleagues' EWAS meta-analysis (Joubert et al., 2016) and their causal effect on 643 complex traits, including phenotypes of depression and anxiety (e.g., *"Seen a psychiatrist for nerves anxiety tension or depression"*) and related phenotypes (e.g., neuroticism). Corresponding cis-SNPs were identified using the mQTL database http://www.mqtldb.org (Gaunt et al., 2016), which is a catalogue of SNP-DNA methylation associations (based on the

450k array) within the ARIES sample (N = 1,018; see Chapter 2 for more details about ARIES) (Relton, Gaunt, et al., 2015). The cis-SNP-outcome associations were obtained from GWAS available on the MR-Base platform http://www.mrbase.org/ (Hemani, Zheng, et al., 2018) and additional GWAS based on UK Biobank data conducted by the Neale lab (http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank). After correction for multiple testing, results of the two-sample MR analysis indicated evidence for one causal association related to mental health. Amongst 22 outcomes that survived correction for multiple testing, DNA methylation at cg01307174 was found to be associated with an increased likelihood of worrying after embarrassment (β = 0.03, SE < 0.01, P-value = 2.65×10^{-08}). Wiklund and colleagues (Wiklund et al., 2019) took a similar approach as Richardson and colleagues (Richardson et al., 2019) by also performing two-sample MR analysis to inspect if there was evidence for causal associations between the differentially methylated CpG sites of prenatal smoking and any of 106 diverse health outcomes (including mental health outcomes) available in the UK biobank. Similar to Richardson and colleagues, they identified cis-SNPs for prenatal-smoking associated DNA methylation from Joubert and colleagues' EWAS meta-analysis (choosing 69 CpG sites with the smallest P-value) and selected corresponding cis-SNPs from the mQTL database (Gaunt et al., 2016). Their analysis found causal evidence for one CpG site (cg25189904) to be associated with the risk for schizophrenia. Except for these two CpG sites, the analyses of Winklund and colleagues and Richardson and colleagues already provide indication that most of the prenatal smoking associated DNA methylation changes are not causally related to the risk of internalising problems.

In this chapter, I follow-up the prenatal smoking-associated CpG sites that were suggestive of contributing to increased worrying and schizophrenia in adulthood. DNA methylation changes at these CpG sites may also contribute to childhood internalising problems, as well as adult anxiety and depression, since these phenotypes are highly correlated with worrying and schizophrenia (Howard et al., 2019; Jami et al., 2020; Nivard et al., 2017; Riglin et al., 2018). Furthermore, both of these studies have not specifically investigated the causal contribution of

smoking associated DNA methylation on *childhood* mental health problems. Due to the large time lag between the prenatal exposure and mental health assessment in the studies of Wiklund and colleagues (Wiklund et al., 2019) and Richardson and colleagues (Richardson et al., 2019), any effect of prenatal smoking-associated DNA methylation on phenotypes of offspring internalising problems may have been diluted by other postnatal influences. Therefore, the analyses of this chapter investigate the association between prenatal smoking- and caffeine-associated DNA methylation signals and early (childhood), as well as later manifestations of internalising problems (adult anxiety and depression).

6.2.2 Genetic of DNA methylation consortium (GoDMC) for cis-SNP selection

Recently, a new data catalogue of cis-SNP-DNA methylation associations has been published. The genetics of DNA methylation consortium (GoDMC) is a large collaboration project between more than 50 research groups that contribute data to investigate the genetic contribution to DNA methylation variation (http://www.godmc.org.uk/) (Min et al., 2020). The GoDMC has investigated the association between genotype and DNA methylation variation using ~10 million genotypes from the 1,000 genome project and blood DNA methylation data from ~30,000 European participants (Min et al., 2020) and is therefore much higher powered than the ARIES based mQTL database (N = 1,018, http://www.mqtldb.org) (Gaunt et al., 2016). The results are now available on a database, containing around 300,000 independent DNA methylation-quantitative trait (mQTL) loci/cis-SNPs (http://mqtldb.godmc.org.uk/). As mentioned above, a cis-SNP represents a SNP that has been associated with DNA methylation at specific CpG site. If the DNA methylation associated SNPs is in proximity to the CpG site (within 1 Mb of each side of the CpG site) it is called a cis-mQTL or cis-SNP (Min et al., 2020). If the mQTL is distal (> 1Mb away from each side of the CpG site) to the CpG site or even on another chromosome, it is commonly referred to trans-mQTL or trans-SNP (Min et al., 2020). Analyses of the GoDMC have confirmed the previous assumption that most mQTLs are cis and only a few are trans (~9%) (Gaunt et al., 2016; Min et al., 2020). In this chapter, I use the GoDMC database for identifying appropriate genetic instruments as proxies for

local DNA methylation associated with prenatal smoking and caffeine exposure on offspring internalising problems using two-sample MR.

6.3 Methods

6.3.1 Data selection

6.3.1.1 Sample 1: Retrieving the cis-SNP – DNA methylation associations

6.3.1.1.1 Smoking

Previous research has found limited evidence for a causal effect of most prenatal-smoking associated DNA methylation on mental health outcomes, except for cg01307174, which showed evidence for a causal effect on worrying, and cg25189904 which showed evidence for a causal effect on schizophrenia (Richardson et al., 2019; Wiklund et al., 2019). Therefore, I selected these CpG sites to investigate their causal effect on internalising problems. The GoDMC identified 790 cis-SNPs to be associated with DNA methylation at cg01307174 and 459 cis-SNPs to be associated with DNA methylation levels at cg25189904.

6.3.1.1.2 Caffeine

In Chapter 4, the probe-level EWAS meta-analysis found evidence for maternal total caffeine consumption to be associated with DNA methylation levels at cg19370043 in cord blood. The GoDMC database identified 5 cis-SNPs to be associated with DNA methylation at cg19370043.

6.3.1.1.3 Internalising problems

In Chapter 5, the probe-level EWAS meta-analysis of childhood internalising problems found evidence for one CpG site (cg26668632) in cord blood to be differentially methylated in association with internalising problems of female offspring at the age of 7. Further, two CpG sites in childhood peripheral blood were found to be differentially methylated in association with childhood internalising problems in female offspring (cg07283896 and cg08884410). The GoDMC showed no cis-SNPs for the CpG site of the cord blood analysis (cg26668632). From the peripheral blood CpG sites, the GoDMC indicated no cis-SNPs for cg08884410 but 74 cis-SNPs to be associated with DNA methylation at cg07283896, which will be used as proxies for childhood internalising problems associated DNA methylation changes.

6.3.1.2 Sample 2: Retrieving the Cis-SNP – outcome association

Overlap between genetic susceptibilities of childhood and adolescent internalising symptoms and adult internalising problems has been reported (Jami et al., 2020). Following the line of reasoning that prenatal exposures may contribute towards the development of internalising problems in childhood and eventually may manifest into adult anxiety and depressive symptoms, I retrieved cis-SNP-outcome associations from the most recent GWAS meta-analysis of childhood internalising problems (Jami et al., 2020), and adult major depression and anxiety (Howard et al., 2019; Purves et al., 2020).

6.3.1.2.1 GWAS of internalising problems in childhood (Jami et al., 2020)

The GWAS meta-analysis of internalising symptoms included repeated data on internalising problems of children aged 3-18 years of age. Data was gathered from 22 cohorts of European ancestry, predominately from the EArly Genetics and Lifecourse Epidemiology (EAGLE) consortium (Middeldorp et al., 2019), resulting in an overall sample size of 64,641 children. Combining data across cohorts generated a GWAS of overall internalising symptoms. In addition to a GWAS analysis of the combined data (N observations = 132,260), cohorts were asked to conduct stratified analysis of age, rater (self, mother, father, teacher), and internalising measure (each with the premise of a sample size of at least 450 children). Stratified cohort results were meta-analysed if a minimum sample size of 15,000 could be reached. Amongst 22 separate measures of internalising problems across cohorts, the same measures as for the EWAS meta-analysis of internalising problems (Chapter 4) were applied by five cohorts (including ALSPAC), contributing data from the emotional subscale of the SDQ and nine cohorts contributing data from the internalising subscale of the CBCL (including MoBa and GenerationR). ALSPAC, GenerationR, and MoBa, which were included in the EWAS meta-analysis of maternal caffeine consumption during pregnancy (Chapter 3) and internalising problems (Chapter 4) were also included in the GWAS meta-analysis of internalising problems, resulting in a sample overlap of ~22%. There was no evidence for any genome-wide significant SNPs in the GWAS meta-analysis of overall internalising problems, nor in any of the stratified analyses. However, there was some evidence of the results capturing some genetic effects of internalising problems. For instance, genetic correlations

showed evidence for association between childhood and adult internalising problems and a polygenic score based on the overall GWAS meta-analysis of internalising problems was able to predict maternal reported internalising problems in children at the age of 7 (explained variance = 0.38%). Furthermore, gene set analyses revealed associations with three genes, including WNT3, which has previously been associated with psychiatric traits related to internalising problems (neuroticism and depressive affect) (Jami et al., 2020).

6.3.1.2.2 GWAS of depression (Howard et al., 2019) and anxiety (Purves et al., 2020) in adulthood

The GWAS meta-analysis of major depression (Howard et al., 2019) included three none-overlapping adult samples (Howard et al., 2018; Hyde et al., 2016; Wray et al., 2018), which led to an overall sample size of 807,553 (246,363 people with depression (including self-declared) and 561,190 without depression). Overall, 102 variants were identified to be associated with depression and most variants showed a similar direction of effect and survived Bonferroni correction (87 variants P-value < 4.90×10^{-4}) or nominal significance (97 variants P-value < 0.05) in an independent replication sample (414,055 people self-reported to have received a clinical diagnosis of depression and 892,299 reported not to have received a diagnosis for depression). A PRS based on the GWAS meta-analysis of depression was able to capture 2% to 3% of variance in depression status (Howard et al., 2019).

The GWAS of anxiety disorders included 126,443 participants of European ancestry from the UK Biobank (25,453 cases anxiety disorder, 58,113 controls; age range = 46-80). Cases of lifetime anxiety disorder were defined as either self-reporting to have received a professional clinical diagnosis for at least one of five anxiety disorders (generalised anxiety disorder, social phobia, panic disorder, agoraphobia, or specific phobia) and/or to meeting a clinical cut-off on the anxiety items of the online administered Composite International Diagnostic Interview (CIDI) Short-form questionnaire (Kessler et al., 1998). Participants with a comorbid self-reported psychiatric diagnosis were excluded from the analysis. In the lifetime anxiety disorder GWAS, five variants were found to be genome-wide significant (P-value < 5×10^{-8}). Further analysis of a genetic instrument generated based on the lifetime anxiety disorder variants was able to predict increased

chances of self-reporting a lifetime anxiety disorder and explained ~1% of variance. Two of the five loci associated with lifetime anxiety disorder were replicated in samples of neuroticism and depression but not in independent, smaller samples of participants with anxiety.

6.3.2 Statistical analysis

6.3.2.1 Enrichment of caffeine and smoking associated CpG sites in the internalising EWAS results

Before conducting MR analyses, preliminary analyses were conducted to explore whether there is initial evidence for the DNA methylation signals from the internalising problems EWAS (Chapter 4) to be enriched for DNA methylation signals of maternal caffeine consumption during pregnancy (Chapter 3) and maternal smoking during pregnancy (Joubert et al., 2016). First, I checked the overlap between the direction of effect of the top 5,000 prenatal smoking- and caffeine-associated CpG sites (ranked by smallest P-value), in the internalising EWAS results from Chapter 4 (if we assume that the prenatal exposure is positively associated with internalising disorders, then we would expect to see the same direction of effect from the prenatal exposure to methylation and from methylation to internalising disorders). Second, I plotted P-values of the associations of the internalising EWAS results at those top 5,000 exposure-associated CpG sites in QQ-Plots. This allowed me to visually inspect whether the associations of the internalising problems EWAS at the top exposure CpG-sites systematically show a smaller P-value than would have been expected by chance.

6.3.2.2 Two-sample MR

The steps for the two-sample MR analyses were as follows: First, I identified cis-SNPs for each exposure through the GoDMC database (http://www.godmc.org.uk/). Second, I extracted summary statistics for the cis-SNPs from the outcome GWAS (childhood internalising problems and adult anxiety and depression). Third, the estimates of the SNP-exposure and SNP-outcome associations were used to generate a causal estimate ($\hat{\beta}_j$). More specifically, the causal estimate was generated for individual variants, using the Wald ratio test (Wald, 1940), by dividing the cis-SNP-outcome estimate ($\widehat{Y}_j$) by the cis-SNP-exposure estimate ($\hat{X}_j$) (Richardson et al., 2019):

$$\hat{\beta}_j = \frac{\hat{Y}_j(cis - SNP_j \rightarrow outcome)}{\hat{X}_j(cis - SNP_j \rightarrow exposure)}$$

If there is more than one cis-SNPs available for a given CpG site, the estimates of the individual variants can be meta-analysed using a fixed-effects inverse-weighted variance approach (Burgess et al., 2013; Richardson et al., 2019):

$$\hat{\beta}_{IVW} = \frac{\sum_j \hat{\beta}_j \hat{X}_j^2 \frac{1}{\sigma_{YK}^2}}{\sum_j \hat{X}_j^2 \frac{1}{\sigma_{YK}^2}}$$

More specifically, the weight reflects the product of the cis-SNP-exposure estimate and the inverse of the variance of the cis-SNP-outcome estimate ($\hat{X}_j^2 \frac{1}{\sigma_{YK}^2}$). The estimated meta-analysed effect is then generated by weighing the individual causal estimates ($\hat{\beta}_{j,\dots i}$) and dividing them by the sum of their weights. The two-sample MR analyses were performed using the *TwoSampleMR* R-package (Hemani, Zheng, et al., 2018), that is linked to the MR-base platform (https://www.mrbase.org/). MR-base is a platform that enables to apply a variety of MR methodologies to curated and harmonised GWAS summary statistics in a user-friendly manner (Hemani, Zheng, et al., 2018). The database includes summary statistics of over 1,673 GWAS and provides all reported SNP-phenotype associations, without restricting to statistically significant results (Hemani, Zheng, et al., 2018). Cis-SNP-DNA methylation associations were extracted using the *read_exposure_data* function. To ensure that the cis-SNPs are independent, clumping was performed using the *clump_data* function using a European reference panel from the 1,000 Genomes Project (Durbin et al., 2010; Hemani, Zheng, et al., 2018). Clumping methods are applied to only keep SNPs that are weakly correlated, while maintaining the SNPs with the strongest signal with a phenotype (see Chapter 2) (Hemani, Zheng, et al., 2018). The function is based on PLINK's clumping method, which prunes SNPs if they show a squared correlation coefficient (r^2) of larger than 0.001 (Hemani, Zheng, et al., 2018). If a cis-SNP was not available in the outcome data, a proxy SNP was searched for, which must show an r^2 with the target SNP of at least 0.80. I retrieved the cis-SNP-outcome association using the *extract_outcome_data* function. Before

running the MR analysis, I harmonised the cis-SNP-exposure and cis-SNP-outcome data using the *harmonise_data* function (Hemani, Zheng, et al., 2018), which ensures that the respective effects were referring to the same allele and effect sizes were on the same scale (e.g., transforming odds rations to β-values using log transformation). Finally, I calculated the two-sample MR causal estimate by running the *mr* function. If only one cis-SNP was available, the function returned the results of the Wald ratio tests for each SNP, and if more than one cis-SNPs was available for use, the function returned the results of a meta-analysis of Wald ratios. The two-sample MR beta coefficients represent the corresponding change in SD of internalising problems/depression/anxiety per one SD change in prenatal smoking-/caffeine-associated DNA methylation.

6.4 Results

6.4.1 Enrichment of smoking and caffeine associated CpG sites in the internalising EWAS results

The look-up of the signals of top prenatal caffeine-CpG sites (Chapter 3) in the cord blood DNA methylation-internalising problems age 3 associations (Chapter 4) showed that ~3% of CpG sites (N CpG sites = 141 of 5,000) had the same direction of effect and P-value < 0.05. This number remained stable for the association of cord blood DNA methylation and internalising problems at the age of 7 (~3%, N CpG sites = 145 of 5,000). Interestingly, the number of prenatal smoking associated CpG sites with the same direction of effect and P-value < 0.05 in the EWAS meta-analysis of internalising problems more than doubled from the age of 3 (~3%, N CpG sites = 128 of 5,000) to the age of 7 (~6%, N CpG sites = 274 of 5,000). This is further supported by QQ-plots of the internalising EWAS P-values just using the top 5,000 prenatal smoking- (Figure 6.4) and caffeine-associated CpG sites (Figure 6.5).

Figure 6.4 reflects what was observed when comparing the direction of effect and P-values of the prenatal smoking- and internalising problems-associated CpG sites. Whereas the P-values of the associations with internalising problems at the age of 3, at the top prenatal smoking-associated CpG sites, show little evidence for enrichment, the associations with internalising problems at the age of 7, at the same CpG sites, show a smaller P-value than would have expected by chance. The

top prenatal smoking CpG site P-values tend to come off the line early on and systematically stay outside the 95% CI. As can be seen in Figure 6.5 below, the P-values of the association of internalising problems in childhood at the top 5,000 prenatal-caffeine associated CpG sites follow the null line closely and stay within the 95% CI, providing little evidence for enrichment.

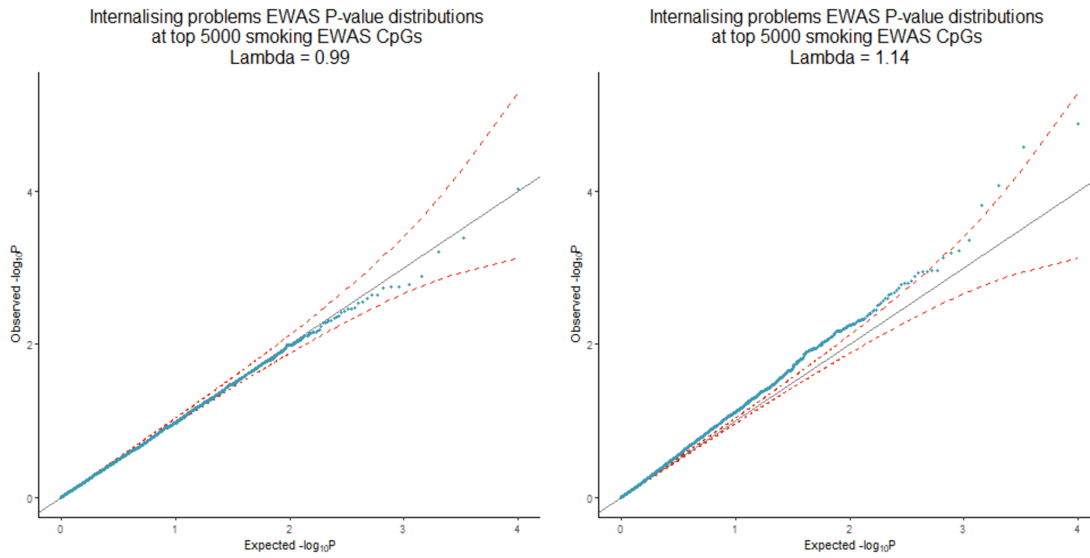


Figure 6.4 QQ-Plots of the association of the EWAS meta-analysis of internalising problems at the top 5,000 prenatal smoking-associated CpG sites. *Association of EWAS meta-analysis of internalising problems at the age of three (left) and the age of seven (right) at the top 5,000 prenatal smoking-associated CpG sites (ranked by smallest P-values). Red dotted lines represent 95% Confidence Intervals (CI). Black line represents H_0 line of P-values not being larger or smaller than expected by chance.*

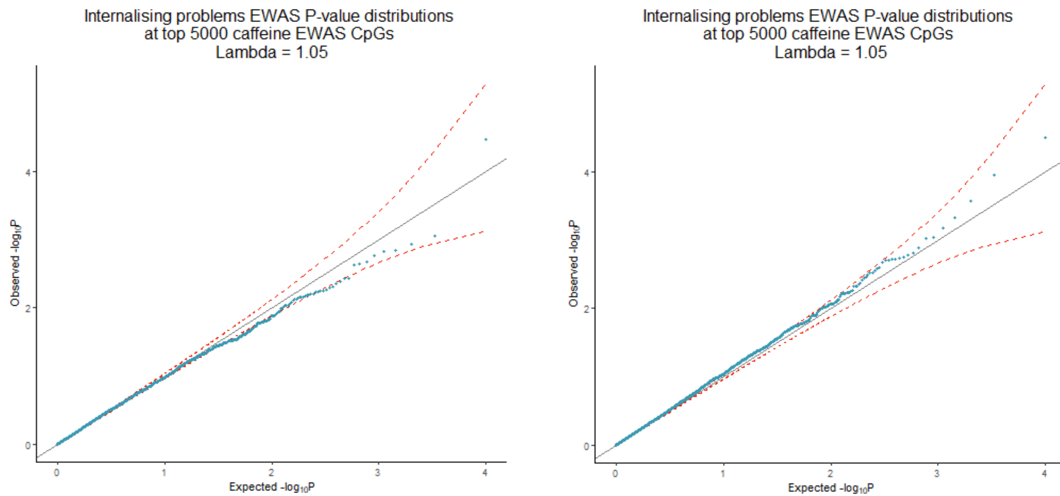


Figure 6.5 QQ-Plots of the association of the EWAS meta-analysis of internalising problems at the top 5,000 prenatal caffeine-associated CpG sites. *Association of EWAS meta-analysis of internalising problems at the age of three (left) and the age of seven (right) at the top 5,000 prenatal caffeine associated CpG sites (ranked by smallest P-values). Red dotted lines represent 95% Confidence Intervals (CI). Black line represents* H_0 *line of P-values not being larger or smaller than expected by chance.*

6.4.2 Clumping of the exposure instruments (cis-SNPs)

For the smoking CpG sites, three independent cis-SNPs remained (rs35062988, rs144982147, rs13085908) after performing clumping on the 790 cis-SNPs of cg01307174. Also, three independent cis-SNPs (rs78434275, rs4655749, rs1276303) of the 459 initially identified cis-SNPs were available after performing clumping on cg25189904. After performing clumping on the five cis-SNPs for prenatal caffeine associated DNA methylation at cg19370043, two independent cis-SNP (rs140512652 and rs139222757) remained, which will be used as a proxy for prenatal caffeine-associated DNA methylation changes in offspring. Performing clumping on the 74 cis-SNPs for childhood internalising problems associated DNA methylation at cg07283896 in peripheral blood in childhood resulted in one independent cis-SNP (rs6930934) that will be used as the corresponding proxy.

6.4.3 Two-sample MR of prenatal smoking-associated DNA methylation and childhood internalising problems, and depression and anxiety in adulthood

For the meta-GWAS of childhood internalising problems, all three cis-SNPs for cg01307174 could be retrieved and two cis-SNPs could be retrieved for cg25189904 (rs4655749, rs1276303). All of the three cis-SNPs for cg01307174 and cg25189904 could be retrieved from the GWAS meta-analysis of adult depression (Howard et al., 2019). For the GWAS of adult anxiety (Purves et al., 2020) rs13085908, and a proxy (rs35458277) for cis-SNP rs35062988, could be retrieved for cg01307174. For DNA methylation at cg25189904 all of the cis-SNPs could be retrieved from the GWAS of adult anxiety (Purves et al., 2020).

The inverse-weighted meta-analysis and Wald ratio tests of the cis-SNPs for DNA methylation at cg01307174 and cg25189904 did not provide evidence for a causal effect of prenatal smoking-associated DNA methylation on either childhood internalising problems or adult depression or anxiety (Table 6.1). For instance, per one SD increase in prenatal smoking associated-DNA methylation at cg01307174, the internalising problems score increased by 0.01 SD (Table 6.1). The amount of DNA methylation variance that the instruments could explain at the two CpG sites associated with prenatal smoking varied from 5%-10% (R^2 exposure in Table 6.1). As expected, the instruments could explain less variance of the outcomes (R^2 outcome, Table 6.1). The SNPs could explain approximately 0.002% to 0.006% of variance of adult anxiety and depression, and 0.0003% to 0.0004% in childhood internalising problems (Table 6.1). This indicates that the direction of effect is as expected, with the instrument influencing smoking-associated DNA methylation first, and then the outcomes through smoking-associated DNA methylation (Hemani et al., 2017).

Table 6.1 Results of the two-sample MR of prenatal smoking-associated DNA methylation changes at birth and childhood internalising problems, and adult anxiety and depression

Outcome	CpG	cis-SNPs available in outcome	Method	Beta (SE)	P-value	R^2 exposure	R^2 outcome
Childhood internalising problems	cg01307174	rs35062988, rs144982147, rs13085908	IVW fixed-effects meta-analysis	0.01 (0.01)	0.587	0.076	2.9×10^{-06}
	cg25189904	rs4655749, rs1276303	IVW fixed-effects meta-analysis	4.7×10^{-03} (0.01)	0.614	0.065	4.0×10^{-06}
Adult depression	cg01307174	rs35062988, rs144982147, rs13085908	IVW fixed-effects meta-analysis	1.5×10^{-03} (0.01)	0.845	0.076	2.14×10^{-05}
	cg25189904	rs4655749, rs1276303	IVW fixed-effects meta-analysis	2.13×10^{-05} (4.1×10^{-03})	0.996	0.098	5.6×10^{-06}
Adult anxiety	cg01307174	rs13085908, rs35458277 (proxy for rs35062988)	IVW fixed-effects meta-analysis	-0.02 (0.18)	0.927	0.043	3.34×10^{-05}
	cg25189904	rs4655749, rs1276303	IVW fixed-effects meta-analysis	0.05 (0.05)	0.334	0.098	1.84×10^{-05}

Note. IVW fixed effects meta-analysis = Inverse-variance weighted fixed-effects meta-analysis.

6.4.4 Two-sample MR of prenatal caffeine-associated DNA methylation and childhood internalising problems, and depression and anxiety in adulthood

Both independent cis-SNPs for prenatal caffeine associated DNA methylation at cg19370043 (rs140512652 and rs139222757) could be retrieved from the GWAS of depression (Howard et al., 2019). None of the two independent cis-SNPs (nor proxy SNPs) could be retrieved from the GWAS of internalising problems (Jami et al., 2020) and the GWAS of adult anxiety (Purves et al., 2020). Therefore, I selected the non-independent cis-SNPs that could be retrieved from the GWAS of internalising problems and adult anxiety. Out of five cis-SNPs for cg19370043, only one cis-SNP could be retrieved from the GWAS of internalising problems (rs17350502) and adult anxiety (rs115822641). The results did not support a causal effect of prenatal caffeine-associated DNA methylation on the risk for internalising problems adult anxiety, or depression (Table 1). The cis-SNPs could explain little variance at the prenatal caffeine-associated CpG site, only accounting for 0.01% to 1.1% of variance (R^2 exposure Table 1), and are therefore considered weak instruments. This was expected, as only single SNPs were available to proxy for DNA methylation at the caffeine-associated CpG site. The variance explained by the exposure cis-SNPs was lower than the explained variance of the outcome cis-SNPs (Table 1, R^2 outcome range: 1.3×10^{-06} to 9.3×10^{-06}), indicating that the causal direction is as expected (Hemani et al., 2017).

Table 6.2 Results of the two-sample MR analysis of prenatal caffeine-associated DNA methylation at birth on childhood internalising problems, and adult anxiety and depression

Outcome	**CpG**	**cis-SNPs available in outcome**	**Method**	**Beta (SE)**	**P-value**	**R^2 exposure**	**R^2 outcome**
Childhood internalising problems	Cg19370043	rs17350502*	Wald ratio	-0.15 (0.16)	0.348	0.0001	8.3×10^{-06}
Adult depression	Cg19370043	rs139222757, rs140512652	IVW fixed-effects meta-analysis	-0.004 (0.01)	0.673	0.011	1.3×10^{-06}
Adult anxiety	Cg19370043	rs115822641*	Wald ratio	0.12 (0.20)	0.545	0.003	9.3×10^{-06}

*Note. * Out of 5 non-independent cis-SNPs for cg1937004 (or proxy SNPs), these were the only SNPs available in the outcome GWAS.*

6.4.5 Two-sample MR of internalising problems-associated DNA methylation and childhood internalising problems, and depression and anxiety in adulthood

The independent cis-SNP, rs6930934, for childhood internalising problems associated DNA methylation at cg07283896 could be retrieved from the GWAS meta-analysis of childhood internalising problems (Jami et al., 2020) and adult depression (Howard et al., 2019). Whereas rs6930934 was not available in the GWAS of anxiety problems (Purves et al., 2020), an appropriate proxy SNP (rs9366903; $R^2 > 0.80$) could be retrieved instead. The Wald Ratio estimates did neither support a causal effect of childhood internalising problems-associated DNA methylation on internalising problems ($\beta < -0.01$, SE = 0.01, P-value = 0.854), nor on adult depression ($\beta = -0.01$, SE = 0.01, P-value = 0.063) or anxiety ($\beta = -0.02$, SE = 0.08, P-value = 0.772). The explained variance of the instruments was approximately 3% (R^2 exposure = 0.031) and the explained variance of the outcomes much smaller at approximately 0.00003% for childhood internalising problems (R^2 exposure = 3.0×10^{-07}), 0.002% for adult depression (R^2 outcome = 1.97×10^{-05}), and 0.00009% for adult anxiety problems (R^2 outcome = 9.0×10^{-07}).

6.5 Discussion

6.5.1 Summary

In this chapter, I investigated if there is causal evidence for an effect of exposure-associated DNA methylation changes at birth on the development of childhood internalising problems, and adult depression and anxiety. The results of the enrichment analysis and two-sample MR analyses consistently provided no evidence for a causal effect of prenatal smoking- and caffeine-associated DNA methylation changes on the risk for any of the outcomes. Furthermore, there was no evidence for childhood internalising problems-associated DNA methylation to be causally related to the onset of internalising problems nor adult depression or anxiety.

Overall, the findings of this chapter align with the results of the other chapters of this thesis: Chapter 3 indicated no evidence for a causal association between maternal smoking and caffeine consumption during pregnancy and offspring

mental health outcomes. Furthermore, Chapters 4 and 5 found limited evidence for an association between prenatal caffeine exposure and internalising problems in childhood with DNA methylation. The QQ-plots of the DNA methylation-internalising problems associations at the top prenatal smoking CpG sites provided some indication for a causal relation between prenatal smoking and internalising problems through DNA methylation. This indicates that the internalising EWAS signals were enriched at prenatal smoking CpG sites despite the internalising EWAS results having been adjusted for maternal smoking (and therefore potentially been biased towards the null). However, a causal effect of prenatal-smoking associated DNA methylation on offspring internalising problems was not supported by the two-sample MR analysis. Testing the effect of DNA methylation at two prenatal smoking associated CpG sites, which in previous research showed evidence for a potential causal contribution towards related phenotypes of internalising problems (increased worrying) (Richardson et al., 2019) and schizophrenia (Wiklund et al., 2019), showed no evidence for a causal effect. Whereas the association between DNA methylation at cg01307174 and internalising problems and depression showed the same direction of effect as with the worrying phenotype in the study by Richardson and colleagues (Richardson et al., 2019), the direction of effect was incongruent for the anxiety results. For DNA methylation at cg25189904, which was found to be causally associated with schizophrenia in the study by Wiklund and colleagues (Wiklund et al., 2019), the direction of effect was reversed for all of the phenotypes of this study. The lack of overlap of the results of this study with the findings of both Richardson's and Wiklund's studies could either indicate that the causal associations of the smoking-associated DNA methylation changes are uniquely contributing to worrying and schizophrenia but not internalising problems, anxiety, or depression, or that these associations represent false positive findings. This former hypothesis is unlikely to hold true, at least for the worrying phenotype, as this is a specific symptom of internalising problems (Wilkinson, 2009). In support for the latter hypothesis is the use of a much higher-powered mqtl database (Min et al., 2020) for identifying appropriate cis-SNPs. Alternatively, it is also possible that stronger instruments, which explain more variance in DNA methylation variation at the selected CpG sites are needed to detect a causal contribution of prenatal smoking- and caffeine-associated DNA

methylation to internalising problems, anxiety, and depression (known as weak instrument bias) (Lawlor, 2016).

6.5.2 Strengths

A major strength of this study is the contribution towards closing the research gap of formal mediation analyses of DNA methylation on developmental mental health outcomes through MR analyses. Whereas the potential for MR analysis to understand the causal contribution of DNA methylation to the development of mental health outcomes have been noted several times (Barker, Walton, & Cecil, 2018; Pingault et al., 2017; Walton et al., 2019), only few studies have applied it (Caramaschi et al., 2018; Cecil et al., 2018). Moreover, the closer investigation of causal associations of prenatal exposures on internalising problems is an important contribution towards the current literature of prenatal exposures on childhood psychopathology, where predominantly the association with externalising problems has been researched (Barker, Walton, & Cecil, 2018). Further, to my knowledge this study was the first to investigate if there was a causal effect of prenatal caffeine-associated DNA methylation on internalising problems (Diemer et al., 2020). Last, the follow-up of two potential causal associations between prenatal smoking associated-DNA methylation and mental health outcomes, using higher-powered cis-SNPs-exposure-associations, provides further evidence against the hypothesis of DNA methylation mediating the effect of prenatal smoking on mental health outcomes in offspring.

6.5.3 Limitations

The study has several limitations. First, the EWAS meta-analysis of prenatal caffeine and internalising problems only found weak evidence for an association with DNA methylation. Thus, it is possible that the DNA methylation levels at the selected exposure-CpG sites were too weakly associated with the exposure to identify an effect on internalising problems. Adding to this, the GWAS meta-analysis of internalising problems also did not find a strong genetic signal with internalising problems (Jami et al., 2020) and thus, despite selecting the best available cis-SNPs through the GoDMC, the two-sample MR may have suffered from noise in the exposure and outcome associations. To combat this, I also conducted the two-sample MR with the adult GWAS of depression (Howard et al., 2018) and anxiety (Purves et al., 2020), in which especially the former was

able to identify many genetic signals. Yet, as these were phenotypes assessed in adulthood, it is possible that too much noise was introduced through the large temporal gap between prenatal exposure and outcome that could have weakened a causal effect. More generally, lack of representation of high smoking/caffeine consumption during pregnancy and offspring with mental health problems may have caused an underestimation of effects for high prenatal smoking/caffeine exposure (see Chapters 1, 2, 3, and 4) (Webb et al., 2017).

Further, it is important to note several limitations of appraising epigenetic changes using MR. First, especially for the caffeine analyses, cis-SNPs were only able to explain a small fraction of variance in DNA methylation and thus, weak instrument bias may have biased the two-sample MR results towards the null (Lawlor, 2016). Second, proper inspection of cis-SNPs as valid instruments was restricted by the limited availability of independent cis-SNPs for the exposure associated CpG sites. In the GoDMC, approximately two cis-SNPs per DNA methylation site of the 450K array have been detected (Min et al., 2020). Few independent cis-SNPs per CpG site restrict the assessment of horizontal pleiotropy and may return spurious associations because of the variants being in LD with a causal variant (Richardson et al., 2018, 2019). Third, as neighbouring CpG sites are believed to exert similar biological functions (Suderman et al., 2018), it may be more appropriate to conduct MR analysis using cis-SNPs that proxy for regional DNA methylation instead of individual CpG sites (Battram et al., 2019). Last, results of the GoDMC indicate that cis-SNPs show low concordance between blood and brain tissue and thus, tissue-specificity may explain the lack of evidence for an effect of changes in blood DNA methylation on internalising problems, anxiety, and depression (Min et al, 2020).

6.5.4 Future directions for epigenetic research in mental health

Recent developments in the field of epigenetics indicate little evidence for a direct contribution of DNA methylation towards complex trait variation (Min et al., 2020). Analyses of the GoDMC indicate that DNA methylation is unlikely to be a direct mediator of genotype-phenotype associations but more likely to influence a variety of dynamic and complex regulatory processes that cause variance in complex traits (Min et al., 2020). This would explain why previous studies also found limited evidence for a causal contribution of DNA methylation to mental

health and other complex traits (Caramaschi et al., 2017, 2018; Cecil et al., 2018; Richardson et al., 2019; Wiklund et al., 2019). Before large scale multi-omics datasets with a larger coverage of DNA methylation probes enable researchers to unravel the potential non-linear paths of DNA methylation to complex trait variation (Min et al., 2020), future research may benefit to focus on developing DNA methylation profiles as biomarkers for mental health problems, and recruiting more diverse samples, which is discussed in more detail in Chapter 7.

6.6 Conclusion

In conclusion, these results, especially in the context of the results of the previous chapters, indicate no evidence for a causal effect of prenatal smoking and caffeine exposure on childhood internalising problems via a mechanism involving DNA methylation. The next chapter discusses these results in the context of the Developmental Origins of Health and Disease framework (DOHaD, see Chapter 1 for more information) and its implications for future research in this area.

Chapter 7 – Discussion

7.1 Chapter overview

The analyses reported in this thesis were intended to contribute to the causal evidence base for the effects of smoking and caffeine consumption during pregnancy on offspring mental health outcomes. As discussed in Chapter 1, causal investigations of prenatal smoking and caffeine exposure are warranted, as the behaviours are quite prevalent in pregnant populations (Anderson et al., 2014; Frary, Johnson, & Wang, 2005; O'Keeffe et al., 2015). Current evidence in support of an effect of smoking and caffeine consumption during pregnancy is predominately based on correlational evidence from observational study designs (e.g., Brion et al., 2010a; Latimer et al., 2012; Miyake et al., 2019; Moylan et al., 2015; Tiesler & Heinrich, 2014). This is problematic because smoking and caffeine consumption during pregnancy are likely to be correlated with other exposures that may also impact offspring's mental health development (such as maternal socioeconomic position, personal hardships, and mental health problems, see Chapter 1). Therefore, correlational observational associations between prenatal smoking and caffeine exposure are likely to be confounded and other methods are needed to establish whether a true causal effect exists. The application of genetic methods and the investigation of DNA methylation as a potential molecular pathway in this thesis complement the observational evidence base. Triangulation of these results helps to unravel which of the reported observational associations between prenatal smoking and caffeine exposure and offspring mental health outcomes represent causal pathways or confounded associations. In this chapter, I summarise the results reported in this thesis and discuss them in the broader context of intrauterine effects, epigenetics, mental health, and the DOHaD framework.

7.2 Summary of thesis findings

In Chapter 3, I conducted a targeted intergenerational PheWAS to explore, in a hypothesis free manner, whether there was evidence for an effect of maternal smoking and caffeine consumption on offspring's mental health phenotypes in childhood. The PRS for smoking and caffeine consumption, which have been

derived in samples from the general population, were validated as proxies for smoking and caffeine consumption during pregnancy. Using maternal and offspring smoking and caffeine PRS as separate predictors for offspring mental health outcomes in childhood, and comparing the magnitude of the corresponding effect estimates, did not show evidence for causal effects of maternal smoking and caffeine consumption on the risk for offspring mental health outcomes. Instead, the comparison of intergenerational (maternal PRS on offspring outcomes in childhood) and childhood effect estimates (offspring PRS on offspring outcomes in childhood) indicated that associations observed in the intergenerational analyses are likely to be attributed to pleiotropic effects. For externalising problems in childhood (hyperactivity, conduct disorders), there was evidence for pleiotropic associations with the smoking genetic variants (Table 7.1). For internalising problems, there was some evidence for pleiotropic associations between the smoking and caffeine genetic variants and a decreased likelihood of anxiety symptoms in childhood. As highlighted in Table 7.2, this analysis was likely underpowered and replication in larger, independent samples is warranted before strong conclusions can be drawn. However, if these results are robust, this could imply that PRS studies using the smoking and caffeine genetic variants to assess the effects of smoking and caffeine on mental health outcomes, may overestimate the effects of mental health phenotypes towards the externalising behaviours spectrum, and underestimate the effects on mental health phenotypes towards the internalising spectrum, especially anxiety phenotypes. The subsequent analysis in this thesis further investigated whether there may be an effect of prenatal smoking and caffeine exposure on offspring internalising problems that may have been underestimated in the PheWAS analysis.

In Chapters 4 and 5, the association between prenatal caffeine and internalising problems in childhood and offspring DNA methylation were investigated through separate EWAS meta-analyses. As highlighted in Table 7.1, both EWAS meta-analyses provided weak evidence for an association between offspring DNA methylation and prenatal exposure to caffeine (Chapter 4) and childhood internalising problems (Chapter 5). This provides little evidence for DNA methylation being a mediator of the relationship between prenatal caffeine exposure and offspring internalising problems. Taken together, the results of

Chapter 3, 4, and 5 already provided lack of evidence for DNA methylation mediating the effect of prenatal caffeine associated DNA methylation on offspring internalising problems. Likewise, for prenatal smoking, the results of the PheWAS analysis in Chapter 3, gave little indication for a causal contribution of prenatal smoking-associated DNA methylation to the risk for offspring mental health problems. This is further supported by previous research reporting lack of evidence for the prenatal smoking-associated DNA methylation changes being causally involved in complex trait variation (Richardson et al., 2019). For completeness, Chapter 6 inspected potential causality of prenatal smoking-and caffeine-associated DNA methylation changes in offspring on the risk for internalising problems, anxiety, and depression. Results indicated no evidence for enrichment of the internalising problems-associated DNA methylation signals at the top 5,000 prenatal caffeine-associated CpG sites but some evidence for enrichment at the top 5,000 prenatal smoking-associated CpG sites. Two-sample MR was conducted using cis-SNPs as proxies for prenatal smoking-, caffeine-, and internalising problems-associated DNA methylation and summary statistics from the largest and most recent GWAS of internalising problems, anxiety and depression (Howard et al., 2019; Jami et al., 2020; Purves et al., 2020). As already expected from the enrichment analyses, as well as the results of the previous chapters, the two-sample MR results provided no evidence for changes in DNA methylation mediating the effect of prenatal smoking, caffeine or internalising problems in childhood on the risk for internalising problems in childhood, and anxiety or depression in adulthood. An overview of the analyses of each chapter and how they contributed towards answering the main research question can be found in Figure 7.1.

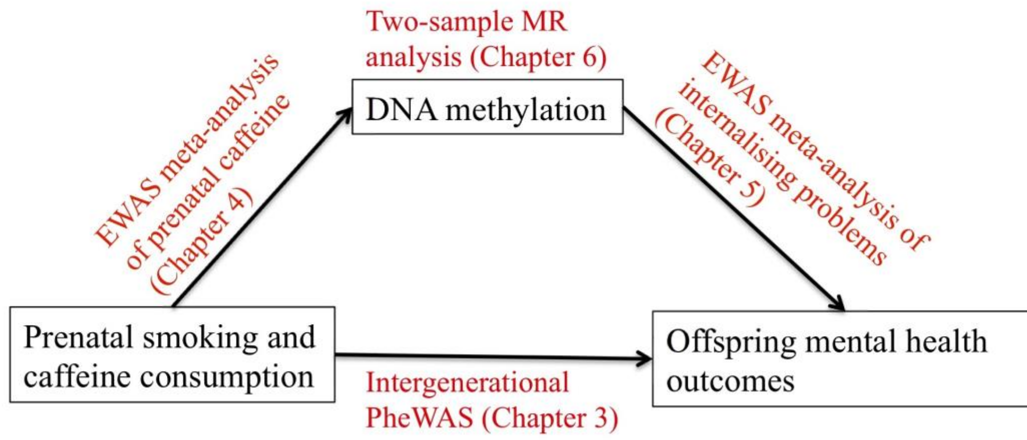


Figure 7.1 Overview of analyses used in this thesis.

Table 7.1 Qualitative Summary of results from each study of this thesis

Chapter	**Main Findings**
3 – Intergenerational PheWAS of prenatal smoking and caffeine polygenic risk scores and offspring mental health outcomes in ALSPAC	- The smoking and caffeine PRS are valid predictors for smoking and caffeine consumption (from tea and coffee but not cola) during pregnancy. - The maternal and offspring smoking PRS showed evidence for associations with an increased risk for childhood externalising problems and decreased anxiety problems. These associations are most likely explained by pleiotropy instead of an effect through maternal smoking. - There was evidence for the maternal caffeine PRS being associated with decreased anxiety problems and sleep during childhood. There was also evidence for the offspring PRS being associated with decreased anxiety during childhood. There is potential for pleiotropy between the caffeine genetic variants and decreased anxiety.
4 – EWAS meta- analysis of prenatal caffeine exposure	- The probe-level and DMR meta-analyses using six European birth cohorts found inconsistent evidence for associations between cord blood DNA methylation and different sources of caffeine (coffee, tea, cola). - Associations that were observed between cord blood DNA methylation and individual sources of caffeine are likely to be explained by other factors than caffeine (e.g., sugar or smoking). - Results provide no evidence for a strong effect of low to moderate amounts of caffeine during pregnancy on offspring DNA methylation.
5 – EWAS meta-analysis of internalising problems in childhood	- Overall, few associations between cord blood and childhood peripheral blood DNA methylation and internalising problems in childhood were found - Both the probe-level and DMR analyses indicated a stronger association with internalising problems in female than in male offspring - Internalising problems amongst the three cohorts included in the meta-analysis were highly zero inflated

	- The DMR analyses suggested differences in DNA methylation at regions annotated to genes that were previously associated with adolescent depression or involved in neurological functioning and development. - The generalizability of these findings should be tested in samples with children with higher expression of internalising problems.
6 – Two-Sample MR of prenatal smoking- and caffeine-associated DNA methylation changes on offspring childhood internalising problems, and adult anxiety and depression	- The QQ-plots of the DNA methylation-internalising problems associations showed some evidence for enrichment at the top 5,000 prenatal smoking-associated CpG sites at the age of 6 (but not at the age of 3). No evidence for enrichment was found at the top 5,000 prenatal caffeine-associated CpG sites. - Two-sample MR analyses did not provide evidence for a causal contribution of DNA methylation at prenatal smoking- and caffeine-associated CpG sites on childhood internalising problems, or anxiety or depression in adulthood.

7.3 Contribution of this thesis towards the existing DOHaD literature

The imbalances of the DoHAD literature outlined in Chapter 1 have been addressed in this thesis as follows. First, this thesis contributes to our understanding of the effects of prenatal caffeine consumption, in addition to the effects of smoking during pregnancy, on offspring mental health outcomes. Smoking and caffeine consumption during pregnancy are highly correlated (Chen et al., 2014; Loomans et al., 2012; Robinson et al., 2010) and have been associated with offspring mental health outcomes (Tiesler & Heinrich, 2014). However, despite these phenotypic correlations, the PheWAS analysis in Chapter 3 did not find evidence for associations between maternal caffeine consumption and offspring mental health outcomes. Also, despite a strong effect of prenatal smoking on offspring cord blood DNA methylation (Joubert et al., 2016) the meta-EWAS of prenatal caffeine exposure in Chapter 4 found limited evidence for an effect of prenatal caffeine exposure on offspring cord blood DNA methylation (Chapter 4). Second, the analyses of this thesis addressed the overrepresentation of studies that have investigated the effect of prenatal smoking on externalising problems, by specifically focussing on internalising problems in Chapters 5 and 6. This was further justified by the findings of the PheWAS analysis in Chapter 3 that indicated that the associations between prenatal smoking and externalising problems in childhood are likely to be explained by pleiotropic effects instead of an intrauterine effect. Third, causal effects were investigated by applying a range of diverse methods to complement the predominant observational evidence base of the DOHaD literature. An overview over the methods used and their corresponding strengths and limitations is presented in Table 7.1. With each of these methods being subject to slightly different biases and still all of the results pointing to the same conclusion, confidence increases in the hypothesis that DNA methylation is not a mediator of prenatal smoking and caffeine consumption on offspring mental health problems. The next section will elaborate on the different underlying biases of each of the analyses conducted within this thesis.

7.3.1 Methodological strengths and limitations

As outlined in Table 7.2, the different methods used in this thesis had unique methodological strengths and limitations, that allowed to inspect the research

question from slightly different angles. Triangulation of the evidence from these different analyses allowed greater confidence in inferences about whether or not an intrauterine effect exists.

The intergenerational PheWAS methodology allowed for a hypothesis free exploration for suggestive causal or pleiotropic associations between smoking and caffeine genetic variants and mental health outcomes. Furthermore, by using PRS as proxies for maternal smoking and caffeine consumption, the number of confounding factors is likely to have been significantly reduced (Davey Smith & Ebrahim, 2003a). The main drawback of this study was the high number of tests that were run, which significantly reduced the statistical power to detect small genetic effects, which are common for complex traits (Choi et al., 2020). Furthermore, despite the PRSs proving to be valid proxies for smoking and caffeine consumption during pregnancy, associations reflect effects of maternal lifetime smoking and caffeine consumption (including pre-and post-pregnancy) and not uniquely intrauterine exposure (Diemer et al., 2020; Lawlor et al., 2017).

The PheWAS analysis was followed by two EWAS meta-analyses to explore the potential of DNA methylation as a molecular pathway linking prenatal caffeine (Chapter 4) and internalising problems in childhood (Chapter 5). A general strength of EWAS analyses is the hypothesis free exploration of DNA methylation in association with traits of interest without limiting analyses to DNA methylation changes at specific genetic regions. Furthermore, dissimilar to the PheWAS analysis, associations are likely to specifically reflect intrauterine exposures, as the EWAS analyses used DNA methylation assessed at birth. However, unlike PRS analyses, results of EWAS analyses are prone to capture associations of confounding variables due to the dynamic nature of DNA methylation. Therefore, I cannot rule out the possibility that the associations observed between prenatal caffeine exposure and DNA methylation are influenced by confounding variables. Also, it needs to be acknowledged that the EWAS analyses in this thesis used DNA methylation assessed by the 450k array and thus results were only able to test associations between 2% of DNA methylation and the phenotypes of interest (Chapter 2). Last, due to tissue specificity the possibility remains that the exposures and internalising problems are associated with DNA methylation changes in brain tissue, that have not been

captured by the analyses in this thesis, which was restricted to DNA methylation assessed in blood.

To get a sense of potential confounding in the EWAS meta-analysis of maternal prenatal caffeine consumption, a negative control and PRS analysis of caffeine consumption during pregnancy was conducted in ALSPAC. Yet, the substantially smaller sample sizes in these analyses, especially in the negative control analysis using paternal self-reported caffeine consumption during pregnancy, hampered the comparison of results with the maternal self-reported caffeine consumption results. Due to statistical power limitations, the negative control and PRS analyses in ALSPAC were unlikely to detect effects of maternal caffeine exposure on offspring DNA methylation, which were indicated to be rather small by the results of the EWAS meta-analysis using maternal self-reported caffeine consumption. More meaningful was the comparison of the associations between DNA methylation and maternal consumption across different caffeinated drinks and different European countries, which are likely to have dissimilar underlying confounding structures. As very few associations were common across the analyses using the different caffeinated drinks, this strengthens the hypothesis that it is not the actual caffeine content that is responsible for associations observed between the caffeinated beverages and offspring DNA methylation (particularly in the higher powered regional DMR analysis). Instead of caffeine, it seems more plausible that the different caffeinated drinks show different confounding structures that influenced DNA methylation (e.g., high sugar content of cola).

In Chapter 6, an enrichment and two-sample MR analysis was applied to investigate the causal relationship between the DNA methylation changes, observed in association with (1) maternal caffeine consumption during pregnancy (Chapter 4), and (2) childhood internalising problems (Chapter 5), on offspring internalising problems. The enrichment analysis had the strength of not restricting test to only the statistically most significant CpG sites but to include the top 5,000 CpG sites with the strongest evidence for associations with prenatal smoking and caffeine consumption. As DNA methylation at neighbouring CpG sites is likely to be highly correlated, this may allow investigating many small effects of exposure-associated DNA methylation that may collectively increase the risk for offspring internalising problems. Yet again, results of the enrichment analysis are only

reflecting EWAS associations, which are prone to confounding and therefore may simply reflect associations with common confounding factors of prenatal smoking and caffeine consumption and offspring internalising problems (e.g., SEP or parental mental health problems). Complementing the enrichment analysis, a two-sample MR analysis was applied that should reduce the likelihood of confounding. A great strength of this approach was the selection of cis-SNPs through the novel and highly powered mqtl database GoDMC (Min et al., 2020) in combination with using two of the most recent and highest powered GWAS of internalising problems (Jami et al., 2020), anxiety (Purves et al., 2020), and depression (Howard et al., 2019). However, the analysis was hampered by the weakly associated signals of local DNA methylation with prenatal caffeine and internalising problems.

A common limitation of all the analyses conducted within this thesis is the presence of selection bias, as reflected by each of the samples showing rather low smoking rates/consuming little caffeine during pregnancy, tending to be more educated, and to suffer from less mental health problems than would be expected in the general population. This limits the generalisability of the findings of this thesis to populations that consume more caffeine or smoke more cigarettes during pregnancy (Webb et al., 2017). Therefore, replication of these analyses is highly warranted in higher risk samples with more extreme manifestations of smoking and caffeine consumption during pregnancy, as well as a higher prevalence of mental health problems in offspring.

Table 7.2 Summary of strengths and limitations of each of the methods used in this thesis

Method	Chapter	Strengths	Limitations
Intergenerational PheWAS	3	- Hypothesis free investigation of smoking and caffeine genetic variants with the mental health "phenome" available in ALSPAC. - Disentangling pleiotropic from potential causal effects of maternal smoking and caffeine consumption on offspring mental health outcomes in childhood.	- Analysis likely to be statistically underpowered because of high number of tests. - Limited to mental health phenotypes assessed in ALSPAC.
PRS analysis	3 and 4	- Reducing confounding variables by using genetic variants for exposures of interest. - Higher predictive power than using single genetic variants (Choi et al., 2020).	- Intergenerational analyses require maternal, paternal, and offspring genetic data to avoid collider bias (Lawlor et al., 2017). - Not specific to intrauterine exposure. PRS for smoking and caffeine is a proxy for lifetime exposure of smoking and caffeine instead exposure during pregnancy (Diemer et al., 2020; Lawlor et al., 2017). - Higher risk of pleiotropic associations because of combining multiple SNPs in one score (Choi et al., 2020). - Requires large sample sizes to detect small genetic effects.
EWAS	4 and 5	- Hypothesis free investigation of associations between DNA methylation and prenatal caffeine and internalising problems. - Insight into potential molecular pathways of pregnancy exposures on offspring mental health outcomes.	- As DNA methylation is dynamic it is subject to confounding to the same degree as the phenome. - Tissue specificity: Unclear whether DNA methylation in cord blood is representative of DNA methylation in most relevant tissue for mental health phenotypes, which is brain tissue. - Only 2% of DNA methylation assessed by 450k array (Bibikova et al., 2011).

Cross-cohort meta-analysis	3, 4, 5, and 6	- Reducing cultural biases because each cohort potentially being subject to slightly different biases that may be cancel each other out in meta-analysed results.	- May introduce heterogeneity due to slightly different measures and assessment time-points across different cohorts and therefore dilute detection of small effects. - Biases shared between cohorts (e.g., selection bias) that may be exacerbated through meta-analysing results.
Negative control analysis using paternal data	3	- Strengthening causal inference of intrauterine effects of caffeine by reducing confounding effects that are shared between maternal and paternal caffeine consumption.	- Highly underpowered due to much smaller sample size of paternal caffeine consumption during pregnancy. -Lack of variability in caffeine consumption because only very few fathers abstained from caffeine consumption during pregnancy.
Enrichment analyses through QQ-plots	6	- Analyses not restricted to few statically significant prenatal smoking and caffeine associated-CpG sites. - Allows to analyse small but systematic associations between internalising problems and smoking- and caffeine- associated CpG sites.	- Associations may reflect effect of common confounding factors of prenatal smoking and caffeine consumption and offspring internalising problems.
Two-Sample MR	6	- Highly powered because of opportunity to use available summary data from publicly available databases. - Reducing confounding variables by using genetic variants as proxies for DNA methylation and outcome of interest.	- Sample overlap may have biased results towards the observational association (Lawlor, 2016). - Dependent on strength of association between DNA methylation and prenatal smoking and caffeine and strength of genetic proxies for exposure (weak instrument bias) of interest (Pierce & Burgess, 2013). - Results may have been biased towards the null because of weak instrument bias (Lawlor, 2016).

7.4 *The future of genetics and epigenetics for prevention and treatment of mental health problems*

The development of reliable biomarkers for mental health phenotypes could contribute to facilitating diagnosis and prognosis of mental health problems, as well as treatment monitoring and clinical decision making (García-Giménez et al., 2017; Ladd-Acosta & Fallin, 2015, Murray 2021). PRS and methylation risk scores (MRS) are promising candidates for biomarker development of mental health problems (Hüls & Czamara, 2020, Murray 2021). Before PRS and MRS biomarkers may be applied to the mental health care setting, more research is needed to increase their predictive power, specificity, and generalizability. The next sections will elaborate on the future of PRS and MRS as biomarkers for mental health problems.

7.4.1 Polygenic risk scores

The rapid developments in the field of genetic epidemiology warrant discussion of how PRS may be applied to clinical psychiatric practice in the future. Since the foundation of the psychiatric genomics consortium in 2009 (https://www.med.unc.edu/pgc/), cheaper and more efficient assessment of genetic information has become available. Collaborative efforts of meta- and mega-analyses enabled recruitment of larger samples and led to a steady increase in predictive power of PRS for mental health problems. Amongst the PRS for psychiatric traits that have been investigated, the PRS for schizophrenia showed the highest predictive power by explaining ~11% of variance, followed by ~4% of variance explained by PRS for bipolar disorder, ADHD, and depression (Murray, 2021). For Autism spectrum disorder and anorexia nervosa, the corresponding PRS were only able to explain ~2% of variance (Murray, 2021). Even though the predictive power of psychiatric traits in not high enough to inform psychiatric practice yet, it is indicated that this may change in the foreseeable future by generating larger sample sizes to increase the variance explained by PRS. For instance, the schizophrenia PRS, which is based on the yet largest GWAS sample size of a psychiatric trait, has approximately the same predictive value as other known risk factors for schizophrenia, such as family history and SES (Agerbo et

al., 2015; Martin et al., 2019). However, merely increasing the sample size of a GWAS did not proof sufficient for obtaining more reliable PRS. For instance, whereas striving for larger sample sizes in GWAS of depression helped to increase the predictive power, it resulted in decreased specificity of the depression PRS. This is likely to be attributed to "minimal phenotyping", where looser definitions of a phenotype are used to save costs and enable inclusion of more participants in the GWAS. However, minimal phenotyping decreases the specificity of genetic variants included in the PRS and therewith compromised the clinical relevance of the PRS (Cai et al., 2020). More detailed assessment of internalising problems may be crucial for genetic and epigenetic research, as the recent GWAS meta-analysis of internalising problems was not successful to identify a signal by merely increasing the sample size (Jami et al., 2020). Future genetic and epigenetic research may benefit from the availability of more precise measures of mental health through the development of digital assessment tools (Ferrar et al., 2020; Goldsack et al., 2020).

With increasing predictive power and specificity of PRS for mental health problems, PRS may become integrated into mental health care in the future. Two recent articles have explored the potential future clinical application of PRS, highlighting that PRS may benefit prevention, diagnosis, and treatment strategies for mental health problems (Wray et al., 2021, Murray et al., 2021). First, PRS may be used to prevent the onset of mental health problems by posing as a screening tool to identify individuals who do not present with symptoms yet, but who might benefit from preventative actions due to an increased genetic risk. For instance, outside of the mental health sector large studies are already being implemented to test the utility of a breast cancer PRS as a screening tool to identify individuals with an increased genetic risk, who may benefit from more frequent monitoring to detect early stages and to engage in protective strategies, such as mastectomies (Wray, 2021; Yanes et al., 2020). There are no current screening strategies applied to identify people at risk for mental illness but if screening programs become implemented in psychiatry, the integration of PRS in such a program may be useful (Murray et al., 2021). The success of commercially attainable direct-to-consumer genetic testing platforms, such as *23andMe* (www.23andme.com), *My Heritage* (www.myheritage.com), and *ancestry*

(www.ancestry.com), are reflecting a general societal interest in genetics and people's willingness to provide their data for genetic screenings. Genetic screenings offered by third party companies are problematic as the methods, which have been applied by these companies to estimate genetic risk, were found to not adhere to scientific standard (e.g., estimating genetic risk based on single SNPs instead of PRS) (Folkersen et al., 2020). To combat this, independent researchers generated a non-profit platform, called *impute.me* (www.impute.me) that enables individuals to receive information about their genetic risk that is in accordance with current scientific standards. The idea of impute.me is that people, who obtained their genetic information through a direct-to-consumer platform, can upload their genetic information to obtain state-of-the-art PRS, that are generated based on the most recent scientific evidence, for more than 2,000 traits (Folkersen et al., 2020). Alarmingly, a survey of 227 impute.me users has found that 61% experienced a negative emotional reaction (upset, anxious, sad, etc.) after receiving their PRS result and that only ~26% of participants could correctly answer questions about PRS and their interpretation (Peck et al., 2021). This stresses the urge for public education and knowledge about genetics to empower people to adequately evaluate and interpret results from such platforms (Folkersen et al., 2020: Murray, 2021). Integrating PRS into health records may be the safest way to ensure that information is appropriately contextualised with other health information and adequately communicated to the patient by a professional. It would also provide an opportunity for the patient to receive appropriate follow-up care through genetic counselling if needed. Second, it has been highlighted that PRS may facilitate clinical diagnosis of mental health disorders by providing another layer of information, especially in instances when individuals present with nonspecific or atypical symptoms that do not clearly fit into a diagnostic category (Murray, 2021). It is important to stress that PRS for complex traits may never be used independently for clinical diagnosis or prevention but may be used as another source of information, in addition to other risk factors such as family history and previous life experiences (Murray, 2021). In the future, if the specificity of PRS improves by identifying disorder specific genetic variants through deeper phenotyping, PRS may also enable a more fine-tuned differentiation of diagnoses (Cai et al., 2020: Murray, 2021). However, this is not possible yet, as many of the currently available mental health PRS are highly

correlated (e.g., correlation of 0.68 between the PRS for schizophrenia and bipolar disorder) (Lee et al., 2013). Last, in addition to screening and diagnoses, future PRS may inform about prognosis and treatment choices for individuals who already have a diagnosis for a mental disorder. For instance, it has been indicated that individuals with a high PRS for schizophrenia are more treatment resistant than individuals with a lower PRS (Murray, 2021). However, current research investigating the utility of PRS for informing about treatment response is limited by small sample sizes due to scarcity of data of individuals that are receiving different forms of treatment (Murray, 2021).

Even more important than increasing the predictive power and specificity of PRS before they become integrated into health care, is the urge to increase the generalizability of PRS to other ancestries than white European ancestries. A review of 3,639 GWAS that were published between 2005 and 2018, investigating 3,508 traits, revealed that, depending on year of publication, 72-96% of the samples came from European ancestry populations (see Figure 7.2) (Mills & Rahal, 2019). This is highly problematic since PRS discovered in European ancestry populations have repeatedly been found to have a much lower predictive power in non-European ancestry populations (Martin et al., 2019). For instance, genetic variants for diabetes that were derived from European ancestry populations were found to lead to false diagnoses in African ancestry populations (Gurdasani et al., 2019; Mills & Tropf, 2020). Therefore, to avoid inflating health disparities, it is crucial that more GWAS are conducted with individuals of more diverse ancestries before PRS may become integrated into health care (Martin et al., 2019).

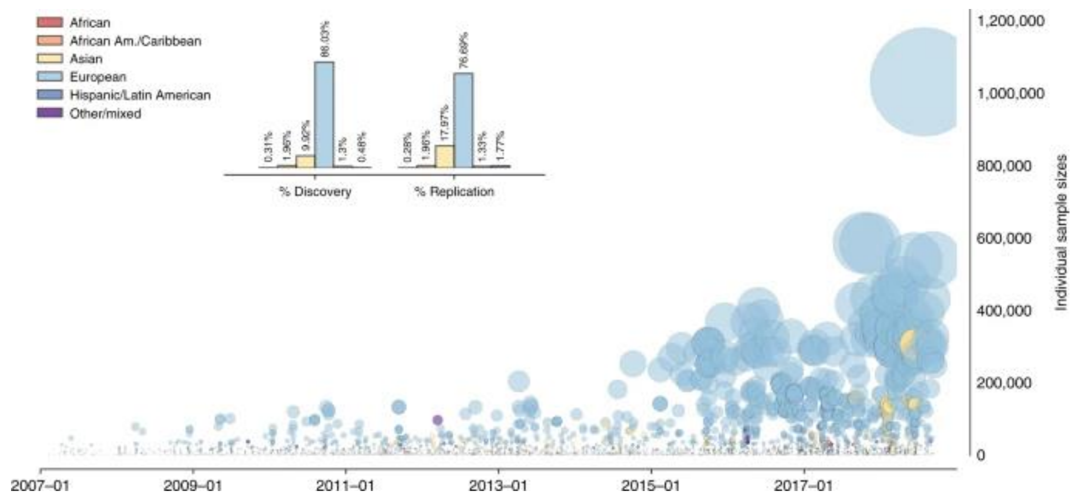


***Figure 7.2 Ancestry populations included in published GWAS** (credit: Mills & Rahal, 2019). X-axis represents date of publication and the y-axis represents sample size of the GWAS. Other/mixed = combinations of ancestries (e.g., European and African). Circles represent sample size per ancestry. The bar graph illustrates the distribution of sample size across different ancestries divide by discovery and replication samples.*

In summary, the genetic epidemiology field is developing quickly and may soon be used to inform clinical psychiatric practice. With larger consortia being established to boost the sample size of GWAS of psychiatric traits, it is important ensure that the specificity of PRS is not sacrificed by relying on single items to assess complex mental health phenotypes. With increasing the predictive power and specificity of PRS for mental health problems, they will become more clinically relevant and may be integrated into health care in the future. Before PRS should be considered as a tool for screening, diagnoses, and treatment, it is of uttermost importance to conduct more GWAS in populations of non-European ancestries to avoid exacerbating health care inequalities. The success of direct-to-consumer genetic testing is warranting a more direct communication and involvement of clinical and research geneticists to ensure that genetic information does not become misinterpreted (Folkersen et al., 2020). Therefore, until PRS become integrated into routine health care it may be important to improve public genetic literacy to facilitate critical and accurate evaluation of genetic testing.

7.4.2 DNA methylation scores

As outlined in the previous chapters, despite theoretical grounds to investigate DNA methylation as an intermediate phenotype of the effect of environmental

exposures on mental health outcomes, scientific evidence to support this theory is lacking. Although DNA methylation at many CpG sites in cord blood was found to be associated with maternal smoking during pregnancy (Joubert et al., 2016), no causal contribution of these DNA methylation changes to the risk of offspring mental health outcomes could be established in Chapter 6. This finding is in line with previous research inspecting the causal contribution of exposure-associated DNA methylation to complex traits (Caramaschi et al., 2018; Cecil et al., 2018; Min et al., 2020; Richardson et al., 2019; Wiklund et al., 2019). This may indicate that, at least for now, it may be beneficial to shift the focus away from the aim of finding causal contribution of DNA methylation to complex trait variation and instead to guide research efforts towards the development of DNA methylation biomarkers. Overall, the lack of causal evidence found in this thesis and previous research indicates that, considering the technology and the accessible tissues (that are not necessarily biological important) available for epidemiological studies, the establishment of DNA methylation biomarkers may be more sensible. Currently association studies of DNA methylation and mental health problems consistently yield only few signals (see Chapter 4) (Barker, Walton, & Cecil, 2018; Emeny et al., 2018; Mooney et al., 2020; Starnawska et al., 2019). Similar as in the initial stages of GWAS of mental health problems (e.g., Purcell et al., 2009), larger samples with DNA methylation data will help detect more signals and enable the development of reliable biomarkers. The development of DNA methylation biomarkers has the advantage of not being reliant on tissue-specificity, even for mental health phenotypes where the most relevant tissue is brain tissue. In fact, DNA methylation biomarkers developed in peripheral tissue, such as blood and saliva, are most desirable as it is easier accessible and less invasive to obtain. However, the establishment of reliable DNA methylation biomarkers requires even larger data sets for training, testing, and validation (Hüls & Czamara, 2020). More DNA methylation data will likely be available in the future due to the rapid development of the field of epigenetics. The exponential increase in epigenetic epidemiology research is represented in Figure 7.3, which shows the rise in publications of epigenetic epidemiology papers as indicated by a search for "epigenetic epidemiology" on PubMed (https://pubmed.ncbi.nlm.nih.gov/). Part of this increase may be explained by a general increase in publications per year. However, this is unlikely to fully explain the very strong increase from 2010, with

less than 300 epigenetic epidemiology publications to more than 800 publications from 2017 onwards.

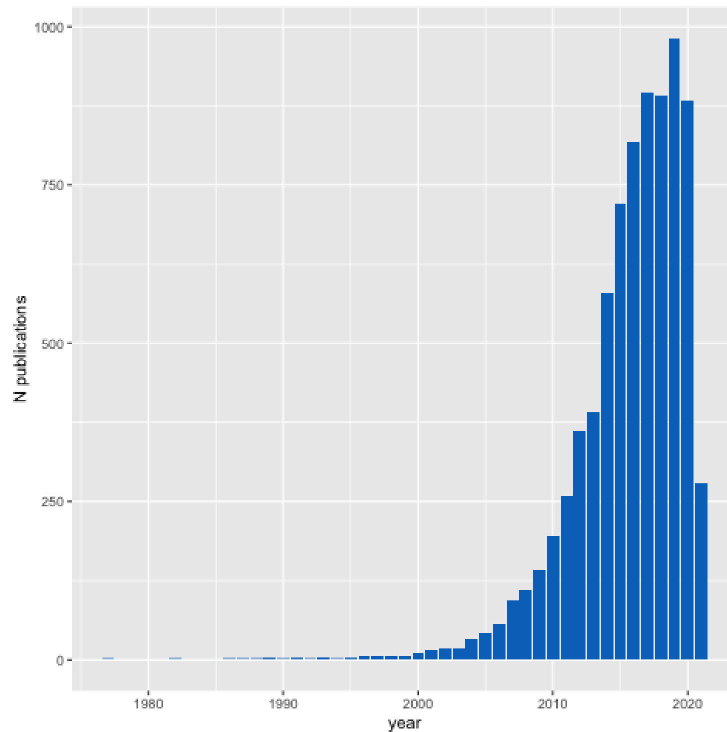


Figure 7.3 Number of epigenetic epidemiology publications per year.

Future directions in the field of epigenetics look promising for developing DNA methylation biomarkers, as more consortia are being established to join forces to detect small epigenetic effects (e.g., PACE) (Felix et al., 2018), GoDMC (Min et al., 2020) and new technologies are being developed that enable capturing a wider spectrum of epigenetic variation in a more cost-efficient manner (Pidsley et al., 2016). Yet, it is important to learn from the pitfalls of GWAS, and to ensure that the specificity of the methylation signals is retained while increasing the sample size of EWAS studies.

Whereas the utility of PRS for clinical prognosis and decision making in psychiatry is already under discussion (see section 7.4.1; Murray et al., 2021; Wray et al., 2021), less literature has explored the utility of methylation risk scores (MRS) for clinical psychiatric practice. Yet, this is a topic worth discussing as recent studies indicate that the generation of MRS may help to establish reliable and highly predictive biomarkers for prenatal exposures and mental health problems in the future. Similar to PRS, MRS are generated by aggregating the many small associations found in association with a phenotype, into one common score that has a higher predictive power (Hüls & Czamara, 2020). The development of MRS requires a large training data set to detect accurate DNA methylation weights, another independent data set the MRS can be tested in and at least one (or more) independent data sets the MRS can be validated in (Hüls & Czamara, 2020). Exploiting the existent DNA methylation data sets, MRS could be developed for a variety of complex traits including smoking and depression (Barbu et al., 2020; McCartney et al., 2018). The MRS for smoking was able to explain 61% of variation in smoking status, compared to 2.8% of variance explained by a PRS in the same sample (McCartney et al., 2018). Together, the MRS and PRS for smoking were able to explain 61.4% of variance in smoking status. Due to its high predictive power and discriminant validity, the smoking MRS has already been recommended to replace phenotypic assessment of smoking in future research (Barbu et al., 2020). Also, biomarkers for maternal smoking during pregnancy in in form of an MRS in offspring cord blood and adult peripheral-blood have been developed (Rauschert et al., 2020; Reese et al., 2017; Richmond et al., 2018). The smoking during pregnancy MRS based on cord blood has shown good accuracy, specificity, and sensitivity for predicting cotinine levels and self-report of sustained smoking during pregnancy (Reese et al., 2017). The exposure to smoking during pregnancy MRS based on offspring's adult peripheral blood (assessed at age 30) showed a lower predictive ability than the MRS based on cord blood but was still able to predict whether mothers smoked during pregnancy (Richmond et al., 2018). Last, another study has applied machine learning methods to develop an MRS for maternal smoking during pregnancy using peripheral blood, which outperformed the MRS from Reese et al.

(2017) and Richmond et al. (2017) based on testing and validating their score in samples of adolescents (age 16) and adults (age 31) (Rauschert et al., 2020).

Whereas MRS may already be used as biomarkers to assess smoking during and outside of pregnancy, more research is needed before appropriate biomarkers for mental health outcomes can be used. To derive good MRS as predictors for phenotypes that show a less strong signal with DNA methylation than smoking, even larger sample sizes will be needed to obtain an appropriate predictive power (Barbu et al., 2020). For instance, a recently developed MRS for depression was only able to explain 1.75% of variance in prevalent and 0.52% in incident depression, which is still lower than the predictive power of the PRS for depression (explained variance = 2.40%). However, the depression MRS was found to contribute a significant amount of explained variance on top of the depression PRS (additive explained variance in depression = 3.99%) (Barbu et al., 2020). Despite more research being needed to improve the predictive power of the depression MRS before it can be applied as a proper biomarker, research using the MRS for depression already yielded valuable insights into the relationship between smoking, depression, and DNA methylation. Compared to the depression MRS trained in a sample of smoking and non-smoking participants, the depression MRS trained solely amongst non-smoking participants showed weaker predictive power for prevalent depression (explained variance = 0.4%) and no clear evidence of association with incident depression (P-value > 0.05). Adding to this, around half of the association between the MRS for depression and risk for depression was mediated by smoking. Another study (Clark et al., 2020) generated MRS from data of 581 major depressive patients and included information from clinical and important lifestyle variables (personality assessment, family history of depression, substance use, etc.) using machine-learning techniques. The MRS, in combination with clinical and lifestyle information, was able to discriminate recurrence of depression within the same individuals 6 years after baseline assessment. This indicates that MRS may inform about disease prognosis and therefore might be useful in the future to guide decisions about appropriate treatments.

It is important to start thinking about the potential utility of MRS for clinical practice. Whereas MRS may be beneficial for clinical research in the near future,

current evidence indicates that more work is needed to improve the sensitivity and specificity of MRS before they can actually be applied to inform clinical decision making (Hillary, 2021). As highlighted in the section on PRS (see 7.4.1), an important step towards developing reliable biomarkers for mental health problems will be to improve the signal to noise ratio in the assessment of mental health phenotypes (Cai et al., 2020). In sum, the development of MRS is an exciting avenue for future epigenetic research and – with larger sample sizes – may help to unravel contributions of environmental exposures on risk for mental health problems in the future. DNA methylation biomarkers could provide a more precise assessment of phenotypes for which self-report may be subject to reporting bias (e.g., smoking during pregnancy, see Chapter 1) or facilitate research in samples where phenotypes of interest have a lot of missing data, or have even not been assessed, by using a MRS as a proxy for those phenotypes (García-Giménez et al., 2017; Ladd-Acosta & Fallin, 2015; Sharp & Relton, 2017).

7.5 Ethical considerations of the development of an epigenetic biosocial archive

Once more reliable MRS are available, they could contribute towards the development of a biosocial archive that encloses information about a person’s life exposures and future health development through epigenetic profiles (Relton, Hartwig, et al., 2015). Whereas the development of a biosocial archive may bring great benefits for understanding intrauterine effects, it also has the potential of inflating already existing health inequalities, if the current underlying biases in the field of DOHaD are not addressed in future epigenetic research. First, epigenetic research must shift away from the strong bias in the DOHaD field of predominantly studying maternal pregnancy exposures (see Chapter 1) (Sharp et al., 2019). As highlighted in Chapter 1, consumption behaviours during pregnancy are embedded in highly complex social structures and are strongly socially patterned. Due to DNA methylation being susceptible to environmental exposures, MRS are likely to capture not only the exposure of interest but also its associated confounding factors. Thus, DNA methylation biomarkers may help to efficiently capture complex confounded structures that would otherwise require a lot of phenotypic assessments (Hillary, 2021; Relton, Hartwig, et al., 2015; Sharp

& Relton, 2017). Yet, this requires that the influences of paternal and societal effects on offspring DNA methylation and mental health outcomes be recognised in addition to any maternal effects. MRS generated based on EWAS of maternal behaviours during pregnancy could act as a proxy for social inequalities that are outside of mothers' actual locus of control. This could lead to mothers being wrongly held solely responsible for their offspring's health outcomes (Ismaili M'hamdi et al., 2018). Therefore, to understand which confounding factors DNA methylation biomarkers are capturing, it is crucial to collect more data on paternal behaviours pre- and post-natally, as well as social factors surrounding increased smoking and caffeine consumption (Easey & Sharp, 2021; Sharp et al., 2019) and to study them in association with offspring DNA methylation. Second, before the development of a biosocial archive efforts should be made to focus collecting epigenetic data from more diverse populations. Current research in the field of epigenetics and complex traits predominantly relies on European ancestry epigenetic samples (such as the PACE consortium and the GoDMC) (Felix et al., 2018; Min et al., 2020). It is very important to widen the representation to other populations, so that the potential benefits of a biosocial archive not only apply to populations of European ancestry. Evidence for the possibility of lack of diverse samples causing health inequality has already been reported from GWAS analyses (see section 7.4.1)(Gurdasani et al., 2019; Mills & Tropf, 2020). As the development of DNA methylation biomarkers is just emerging, it is very important to immediately address these imbalances in epigenetic research investigating DOHaD hypotheses to prevent health inequalities to be amplified by the emerging biosocial archive. Current large epigenetic consortia, such as the PACE (Felix et al., 2018) and GoDMC consortium (Min et al., 2020) mostly rely on white European/US populations, making it difficult to study DNA methylation in populations from diverse ethnicities. However, some of the more recent cohorts started to put effort into sampling from more diverse populations. For instance, the Born in Bradford cohort, which was included in the EWAS meta-analysis of prenatal caffeine of this thesis (Chapter 4), invested great efforts into recruiting participants from South-Asian ethnicities. Strategies applied in the recruitment of this cohort, such as translating questionnaires into different languages and providing translators for the administration of questionnaires (Chapter 2), may help future cohorts to increase their ethnic diversity.

7.6 Main conclusion

Overall, the genetically and epigenetically informed analyses of this thesis, that each had slightly varying underlying biases, all provided no evidence for a causal effect of prenatal smoking and caffeine exposure on offspring mental health problems, including via a mechanism involving DNA methylation. The associations between the smoking PRS and offspring externalising problems found in Chapter 3 showed more evidence for being confounded by pleiotropy instead of reflecting a causal effect of prenatal smoking on offspring externalising problems. Likewise, the lacking synergy between associations of DNA methylation and consumption of different caffeinated drinks during pregnancy found in Chapter 4, indicates that different confounding structures of different caffeinated drinks are more likely to explain the associations than a causal intrauterine effect of caffeine on offspring DNA methylation. These findings were supported by the results of Chapter 6, which did not find evidence for a causal effect of prenatal smoking- and caffeine-associated DNA methylation on internalising problems, anxiety, and depression.

Whereas the results of my thesis should not be used independently to advise expectant parents about intrauterine effects of smoking and caffeine on mental health outcomes, they may be triangulated with other research findings, as well as guide directions for further research in the field of DOHaD. For instance, the finding of potential pleiotropy of smoking and externalising behaviour traits is supported by other research using different methodologies and may be used to enhance prevention and intervention programs for stopping smoking during pregnancy (Crone & Reijneveld, 2007; Elkins et al., 2007; Pedersen et al., 2018; Harrison et al., 2019; Hicks et al., 2020). Whereas abstaining from smoking during pregnancy is highly recommended, due to the established harmful effects on offspring health, more research is needed to increase understanding of the effects of intrauterine caffeine exposure on offspring health. The results of my thesis have shown that investigating different sources of caffeine may help to differentiate confounded from causal effects of caffeine. The lack of evidence for a causal effect of DNA methylation at birth on mental health phenotypes is also in line with recent findings. Results of the highly statistically powered collaborative effort of the GoDMC have emphasised that – given the currently available

technologies – it is unlikely that any causal contribution of DNA methylation to complex trait variation may be discovered (Min et al, 2020). Before better technologies enable to assess a larger fraction of DNA methylation markers in more diverse tissue and cell types, DNA methylation may be used as a potential biomarker that could replace self-report measures and pose as a more accurate assessment of intrauterine exposure to smoking and caffeine consumption and mental health problems. If genetic and epigenetic research manages to address the current biases underlying the GWAS, EWAS, and DOHaD literature, PRS and MRS for mental health outcomes and intrauterine exposures could contribute towards improving prevention, diagnoses, and treatment of offspring mental health problems through the development of a biosocial archive. To do so, more diverse samples need to be recruited, which represent more diverse ancestries and risk populations. Furthermore, more data needs to be collected from fathers (not just mothers) and the wider social context (e.g., social inequalities) that is influencing human behaviour.

Considering the currently lacking scientific evidence for a causal intrauterine effect of caffeine on offspring health problems, polarized media headlines and strong public opinions about consumption of caffeine during pregnancy appear disproportionate. The biases of the DOHaD field that have been described in Chapter 1, stress the importance of research findings being transparently communicated without using inflammatory language and deriving premature advice for pregnant women from single studies. A collaborative effort between the British Pregnancy Advisory Service and researcher from Cardiff University formed the WRISK project that aims at enhancing risk communication for pregnancy by advocating for a women-centred perspective when communicating pregnancy related risks (www.wrisk.org). A study as part of the WRISK project found that amongst 171 randomly selected media headlines published in the UK between 2018 and 2019, the most talked about risk factors were related to maternal food and drink consumption during pregnancy (Marshall et al., 2021). In the sample of headlines that they investigated, headlines were predominantly focussed on fetal outcomes, ignoring effects on maternal health. Interestingly, no strong deviations from the content of media headlines and studies' press releases could be identified, stressing the importance of adequate and sensitive press

releases and appropriate communication from the Science Media Centres of the corresponding research institutions. I hope that future work in DOHaD will foster a more considerate research culture, which openly acknowledges the complexity of investigating causal intrauterine effects, while empowering women to make an informed choice when facing uncertainty about the effects of behaviours during pregnancy.

References

Abbott, L. C., & Winzer-Serhan, U. H. (2012). Smoking during pregnancy: Lessons learned from epidemiological studies and experimental studies using animal models. *Critical Reviews in Toxicology*, *42*(4), 279–303. https://doi.org/10.3109/10408444.2012.658506

Abdel-Hady, H., Nasef, N., Shabaan, A. E., & Nour, I. (2015). Caffeine therapy in preterm infants. *World Journal of Clinical Pediatrics*, *4*(4), 81–93. https://doi.org/10.5409/wjcp.v4.i4.81

Abdul-Hussein, A., Kareem, A., Tewari, S., Bergeron, J., Briollais, L., Challis, J. R., Davidge, S. T., Delrieux, C., Fortier, I., Goldowitz, D., & others. (2020). Early life risk and resiliency factors and their influences on developmental outcomes and disease pathways: A rapid evidence review of systematic reviews and meta-analyses. *Journal of Developmental Origins of Health and Disease*, 1–16.

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs*, *80*(7), 1–37. WorldCat.org.

Achenbach, T. M., Ivanova, M. Y., Rescorla, L. A., Turner, L. V., & Althoff, R. R. (2016). Internalizing/Externalizing Problems: Review and Recommendations for Clinical and Research Applications. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(8), 647–656. https://doi.org/10.1016/j.jaac.2016.05.012

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles* (Vol. 30). Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Agerbo, E., Sullivan, P. F., Vilhjálmsson, B. J., Pedersen, C. B., Mors, O., Børglum, A. D., Hougaard, D. M., Hollegaard, M. V., Meier, S., Mattheisen, M., Ripke, S., Wray, N. R., & Mortensen, P. B. (2015). Polygenic Risk Score, Parental Socioeconomic Status, Family History of Psychiatric Disorders, and the Risk for Schizophrenia: A Danish Population-Based Study and Meta-analysis. *JAMA Psychiatry*, *72*(7), 635. https://doi.org/10.1001/jamapsychiatry.2015.0346

Ahun, M. N., Consoli, A., Pingault, J.-B., Falissard, B., Battaglia, M., Boivin, M.,

Tremblay, R. E., & Côté, S. M. (2018). Maternal depression symptoms and internalising problems in the offspring: The role of maternal and family factors. *European Child & Adolescent Psychiatry*, *27*(7), 921–932.

Aktar, E., Qu, J., Lawrence, P. J., Tollenaar, M. S., Elzinga, B. M., & Bögels, S. M. (2019). Fetal and Infant Outcomes in the Offspring of Parents With Perinatal Mental Disorders: Earliest Influences. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/fpsyt.2019.00391

Alasmari, F. (2020). Caffeine induces neurobehavioral effects through modulating neurotransmitters. *Saudi Pharmaceutical Journal*, *28*(4), 445–451. https://doi.org/10.1016/j.jsps.2020.02.005

Alati, R., MacLeod, J., Hickman, M., Sayal, K., May, M., Smith, G. D., & Lawlor, D. A. (2008). Intrauterine Exposure to Alcohol and Tobacco Use and Childhood IQ: Findings from a Parental-Offspring Comparison within the Avon Longitudinal Study of Parents and Children. *Pediatric Research*, *64*(6), 659–666. https://doi.org/10.1203/PDR.0b013e318187cc31

Alfano, R., Guida, F., Galobardes, B., Chadeau-Hyam, M., Delpierre, C., Ghantous, A., Henderson, J., Herceg, Z., Jain, P., Nawrot, T. S., Relton, C., Vineis, P., Castagné, R., & Plusquin, M. (2019). Socioeconomic position during pregnancy and DNA methylation signatures at three stages across early life: Epigenome-wide association studies in the ALSPAC birth cohort. *International Journal of Epidemiology*, *48*(1), 30–44. https://doi.org/10.1093/ije/dyy259

Ali-Khan, S. E., Krakowski, T., Tahir, R., & Daar, A. S. (2011). The use of race, ethnicity and ancestry in human genetic research. *The HUGO Journal*, *5*(1–4), 47–63. https://doi.org/10.1007/s11568-011-9154-5

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nature Protocols*, *5*(9), 1564–1573. https://doi.org/10.1038/nprot.2010.116

Andre, Q. R., Geeraert, B. L., & Lebel, C. (2020). Brain structure and internalizing and externalizing behavior in typically developing children and adolescents. *Brain Structure and Function*, *225*(4), 1369–1378.

https://doi.org/10.1007/s00429-019-01973-y

Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, *5*(4), 237–249.

Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry*, *40*(1), 57–87. https://doi.org/10.1111/1469-7610.00424

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, *30*(10), 1363–1369.

Ashford, J., Lier, P. A. C. V., Timmermans, M., Cuijpers, P., & Koot, H. M. (2008). Prenatal Smoking and Internalizing and Externalizing Problems in Children Studied From Childhood to Late Adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(7), 779–787. https://doi.org/10.1097/CHI.0b013e318172eefb

Audrain-McGovern, J., Rodriguez, D., Tercyak, K. P., Cuevas, J., Rodgers, K., & Patterson, F. (2004). Identifying and Characterizing Adolescent Smoking Trajectories. *Cancer Epidemiol Biomarkers Prev*, 13.

Bai, J., Wong, F. W., Bauman, A., & Mohsin, M. (2002). Parity and pregnancy outcomes. *American Journal of Obstetrics and Gynecology*, *186*(2), 274–278.

Ballering, G., Leijnse, J., Eijkelkamp, N., Peeters, L., & Heus, R. de. (2018). First-trimester placental vascular development in multiparous women differs from that in nulliparous women. *The Journal of Maternal-Fetal & Neonatal Medicine*, *31*(2), 209–215. https://doi.org/10.1080/14767058.2017.1280020

Bakker, R., Steegers, E. A., Obradov, A., Raat, H., Hofman, A., & Jaddoe, V. W. (2010). Maternal caffeine intake from coffee and tea, fetal growth, and the risks of adverse birth outcomes: The Generation R Study. *The American Journal of Clinical Nutrition*, *91*(6), 1691–1698. https://doi.org/10.3945/ajcn.2009.28792

Bakulski, K. M., Feinberg, J. I., Andrews, S. V., Yang, J., Brown, S., L.

McKenney, S., Witter, F., Walston, J., Feinberg, A. P., & Fallin, M. D. (2016). DNA methylation of cord blood cell types: Applications for mixed cell birth studies. *Epigenetics*, *11*(5), 354–362. https://doi.org/10.1080/15592294.2016.1161875

Baler, R. D., Volkow, N. D., Fowler, J. S., & Benveniste, H. (2008). Is fetal brain monoamine oxidase inhibition the missing link between maternal smoking and conduct disorders? *Journal of Psychiatry & Neuroscience : JPN*, *33*(3), 187–195.

Barbu, M. C., Shen, X., Walker, R. M., Howard, D. M., Evans, K. L., Whalley, H. C., Porteous, D. J., Morris, S. W., Deary, I. J., Zeng, Y., Marioni, R. E., Clarke, T.-K., & McIntosh, A. M. (2020). Epigenetic prediction of major depressive disorder. *Molecular Psychiatry*, 1–12. https://doi.org/10.1038/s41380-020-0808-3

Barker, E. D., Copeland, W., Maughan, B., Jaffee, S. R., & Uher, R. (2012). The Relative Impact of Maternal Depression and Associated Risk Factors on Offspring Psychopathology. *The British Journal of Psychiatry : The Journal of Mental Science*, *200*(2), 124–129. https://doi.org/10.1192/bjp.bp.111.092346

Barker, E. D., Walton, E., & Cecil, C. A. M. (2018). Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, *59*(4), 303–322. https://doi.org/10.1111/jcpp.12782

Barker, E. D., Walton, E., Cecil, C. A. M., Rowe, R., Jaffee, S. R., Maughan, B., O’Connor, T. G., Stringaris, A., Meehan, A. J., McArdle, W., Relton, C. L., & Gaunt, T. R. (2018). A Methylome-Wide Association Study of Trajectories of Oppositional Defiant Behaviors and Biological Overlap With Attention Deficit Hyperactivity Disorder. *Child Development*, *89*(5), 1839–1855. https://doi.org/10.1111/cdev.12957

Barnholtz-Sloan, J. S., McEvoy, B., Shriver, M. D., & Rebbeck, T. R. (2008). Ancestry Estimation and Correction for Population Stratification in Molecular Epidemiologic Association Studies. *Cancer Epidemiology Biomarkers & Prevention*, *17*(3), 471–477. https://doi.org/10.1158/1055-9965.EPI-07-0491

Battram, T., Richmond, R. C., Baglietto, L., Haycock, P. C., Perduca, V.,

Bojesen, S. E., Gaunt, T. R., Hemani, G., Guida, F., Carreras-Torres, R., Hung, R., Amos, C. I., Freeman, J. R., Sandanger, T. M., Nøst, T. H., Nordestgaard, B. G., Teschendorff, A. E., Polidoro, S., Vineis, P., … Relton, C. L. (2019). Appraising the causal relevance of DNA methylation for risk of lung cancer. *International Journal of Epidemiology*, *48*(5), 1493–1504. https://doi.org/10.1093/ije/dyz190

Bech, B. H., Obel, C., Henriksen, T. B., & Olsen, J. (2007). Effect of reducing caffeine intake on birth weight and length of gestation: Randomised controlled trial. *BMJ : British Medical Journal*, *334*(7590), 409. https://doi.org/10.1136/bmj.39062.520648.BE

Beck, C. T. (2001). Predictors of postpartum depression: An update. *Nursing Research*, *50*(5), 275–285.

Bekkhus, M., Skjøthaug, T., Nordhagen, R., & Borge, A. (2010). Intrauterine exposure to caffeine and inattention/overactivity in children: Caffeine and inattention/overactivity. *Acta Paediatrica*, *99*(6), 925–928. https://doi.org/10.1111/j.1651-2227.2010.01744.x

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., & Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behavioural Brain Research*, *125*(1), 279–284. https://doi.org/10.1016/S0166-4328(01)00297-2

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berglundh, S., Vollrath, M., Brantsæter, A. L., Brandlistuen, R., Solé-Navais, P., Jacobsson, B., & Sengpiel, V. (2020). Maternal caffeine intake during pregnancy and child neurodevelopment up to eight years of age—Results from the Norwegian Mother, Father and Child Cohort Study. *European Journal of Nutrition*. https://doi.org/10.1007/s00394-020-02280-7

Bibikova, M. (2006). High-throughput DNA methylation profiling using universal bead arrays. *Genome Research*, *16*(3), 383–393. https://doi.org/10.1101/gr.4410706

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., Delano, D., Zhang, L., Schroth, G. P., Gunderson, K. L., Fan, J.-B., & Shen, R. (2011).

High density DNA methylation array with single CpG site resolution. *Genomics*, *98*(4), 288–295. https://doi.org/10.1016/j.ygeno.2011.07.007

Bird, A. (2007). Perceptions of epigenetics. *Nature*, *447*(7143), 396–398. https://doi.org/10.1038/nature05913

Bjørngaard, J. H., Nordestgaard, A. T., Taylor, A. E., Treur, J. L., Gabrielsen, M. E., Munafò, M. R., Nordestgaard, B. G., Åsvold, B. O., Romundstad, P., & Davey Smith, G. (2017). Heavier smoking increases coffee consumption: Findings from a Mendelian randomization analysis. *International Journal of Epidemiology*, *46*(6), 1958–1967. https://doi.org/10.1093/ije/dyx147

Blood-Siegfried, J., & Rende, E. K. (2010). The Long-Term Effects of Perinatal Nicotine Exposure on Neurologic Development. Journal of Midwifery & Women's Health, 55(2), 143–152. https://doi.org/10.1016/j.jmwh.2009.05.006

Borenstein, M., Hedges, L., & Rothstein, H. (2007). Meta-analysis: Fixed effect vs. Random effects. *Meta-Analysis. Com*.

Bowden, J., Davey Smith, G., & Burgess, S. (2015). Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *International Journal of Epidemiology*, *44*(2), 512–525. https://doi.org/10.1093/ije/dyv080

Bowker, K., Lewis, S., Coleman, T., & Cooper, S. (2015). Changes in the rate of nicotine metabolism across pregnancy: A longitudinal study. *Addiction (Abingdon, England)*, *110*(11), 1827–1832. https://doi.org/10.1111/add.13029

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., & Davey Smith, G. (2013). Cohort Profile: The 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, *42*(1), 111–127. https://doi.org/10.1093/ije/dys064

Boylan, S. M., Cade, J. E., Kirk, S. F. L., Greenwood, D. C., White, K. L. M., Shires, S., Simpson, N. A. B., Wild, C. P., & Hay, A. W. M. (2008). Assessing caffeine exposure in pregnant women. *The British Journal of Nutrition*, *100*(4), 875–882. https://doi.org/10.1017/S0007114508939842

Braun, P. R., Han, S., Hing, B., Nagahama, Y., Gaul, L. N., Heinzman, J. T.,

Grossbach, A. J., Close, L., Dlouhy, B. J., Howard, M. A., Kawasaki, H., Potash, J. B., & Shinozaki, G. (2019). Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. *Translational Psychiatry*, *9*(1), 1–10. https://doi.org/10.1038/s41398-019-0376-y

Brion, M.-J., Victora, C., Matijasevich, A., Horta, B., Anselmi, L., Steer, C., Menezes, A. M. B., Lawlor, D. A., & Smith, G. D. (2010). Maternal smoking and child psychological problems: Disentangling causal and non-causal effects. *Pediatrics*, *126*(1), e57–e65. https://doi.org/10.1542/peds.2009-2754

Brummelte, S., Mc Glanaghy, E., Bonnin, A., & Oberlander, T. (2017). Developmental changes in serotonin signaling: Implications for early brain function, behavior and adaptation. *Neuroscience*, *342*, 212–231.

Bubier, J. L., & Drabick, D. A. G. (2009). Co-occurring Anxiety and Disruptive Behavior Disorders: The Roles of Anxious Symptoms, Reactive Aggression, and Shared Risk Processes. *Clinical Psychology Review*, *29*(7), 658–669. https://doi.org/10.1016/j.cpr.2009.08.005

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, P. L., Amode, R., Guillen, J. A., Riat, H. S., Trevanion, S. J., Hall, P., Junkins, H., … Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, *47*(D1), D1005–D1012. https://doi.org/10.1093/nar/gky1120

Burgess, S., Butterworth, A., & Thompson, S. G. (2013). Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. *Genetic Epidemiology*, *37*(7), 658–665. https://doi.org/10.1002/gepi.21758

Burgess, S., Thompson, S. G., & CRP CHD Genetics Collaboration. (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *International Journal of Epidemiology*, *40*(3), 755–764. https://doi.org/10.1093/ije/dyr036

Burt, S. A. (2009). Rethinking Environmental Contributions to Child and Adolescent Psychopathology: A Meta-Analysis of Shared Environmental

Influences. *Psychological Bulletin*, *135*, 608–637. https://doi.org/10.1037/a0015702

Buscariollo, D. L., Fang, X., Greenwood, V., Xue, H., Rivkees, S. A., & Wendler, C. C. (2014). Embryonic Caffeine Exposure Acts via A1 Adenosine Receptors to Alter Adult Cardiac Function and DNA Methylation in Mice. *PLoS ONE*, *9*(1), e87547. https://doi.org/10.1371/journal.pone.0087547

Bush, W. S., Oetjens, M. T., & Crawford, D. C. (2016). Unravelling the human genome–phenome relationship using phenome-wide association studies. *Nature Reviews Genetics*, *17*(3), 129–145. https://doi.org/10.1038/nrg.2015.36

Cai, N., Revez, J. A., Adams, M. J., Andlauer, T. F. M., Breen, G., Byrne, E. M., Clarke, T.-K., Forstner, A. J., Grabe, H. J., Hamilton, S. P., Levinson, D. F., Lewis, C. M., Lewis, G., Martin, N. G., Milaneschi, Y., Mors, O., Müller-Myhsok, B., Penninx, B. W. J. H., Perlis, R. H., … Flint, J. (2020). Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nature Genetics*, *52*(4), 437–447. https://doi.org/10.1038/s41588-020-0594-5

Calker, D. van, Biber, K., Domschke, K., & Serchov, T. (2019). The role of adenosine receptors in mood and anxiety disorders. *Journal of Neurochemistry*, *151*(1), 11–27. https://doi.org/10.1111/jnc.14841

Cao-Lei, L., de Rooij, S. R., King, S., Matthews, S. G., Metz, G. A. S., Roseboom, T. J., & Szyf, M. (2017). Prenatal stress and epigenetics. *Neuroscience & Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2017.05.016

Caramaschi, D., Hatcher, C., Mulder, R. H., Felix, J. F., Cecil, C. A. M., Relton, C. L., & Walton, E. (2020). Epigenome-wide association study of seizures in childhood and adolescence. *Clinical Epigenetics*, *12*(1), 8. https://doi.org/10.1186/s13148-019-0793-z

Caramaschi, D., Sharp, G. C., Nohr, E. A., Berryman, K., Lewis, S. J., Davey Smith, G., & Relton, C. L. (2017). Exploring a causal role of DNA methylation in the relationship between maternal vitamin B12 during pregnancy and child's IQ at age 8, cognitive performance and educational attainment: A two-step Mendelian randomization study. *Human Molecular Genetics*, *26*(15), 3001–3013. https://doi.org/10.1093/hmg/ddx164

Caramaschi, D., Taylor, A. E., Richmond, R. C., Havdahl, K. A., Golding, J., Relton, C. L., Munafò, M. R., Davey Smith, G., & Rai, D. (2018). Maternal smoking during pregnancy and autism: Using causal inference methods in a birth cohort study. *Translational Psychiatry*, *8*(1), 1–10. https://doi.org/10.1038/s41398-018-0313-5

CARE Study Group. (2008). Maternal caffeine intake during pregnancy and risk of fetal growth restriction: A large prospective observational study. *BMJ*, *337*(nov03 2), a2332–a2332. https://doi.org/10.1136/bmj.a2332

Carslake, D., Tynelius, P., van den Berg, G., Davey Smith, G., & Rasmussen, F. (2017). Associations of parental age with health and social factors in adult offspring. Methodological pitfalls and possibilities. *Scientific Reports*, *7*(1), 45278. https://doi.org/10.1038/srep45278

Cecil, C. A. M., Walton, E., Jaffee, S. R., O'Connor, T., Maughan, B., Relton, C. L., Smith, R. G., McArdle, W., Gaunt, T. R., Ouellet-Morin, I., & Barker, E. D. (2018). Neonatal DNA methylation and early-onset conduct problems: A genome-wide, prospective study. *Development and Psychopathology*, *30*(2), 383–397. https://doi.org/10.1017/S095457941700092X

Chamberlain, C., O'Mara-Eves, A., Porter, J., Coleman, T., Perlen, S. M., Thomas, J., & McKenzie, J. E. (2017). Psychosocial interventions for supporting women to stop smoking in pregnancy. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD001055.pub5

Chen, L., Bell, E. M., Browne, M. L., Druschel, C. M., & Romitti, P. A. (2014). Exploring Maternal Patterns of Dietary Caffeine Consumption Before Conception and During Pregnancy. *Maternal and Child Health Journal*, *18*(10), 2446–2455. https://doi.org/10.1007/s10995-014-1483-2

Chen, Y. A., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., Gallinger, S., Hudson, T. J., & Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*, *8*(2), 203–209. https://doi.org/10.4161/epi.23470

Chen, Y., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., Gallinger, S., Hudson, T. J., & Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium

HumanMethylation450 microarray. *Epigenetics : Official Journal of the DNA Methylation Society*, *8*(2), 203–209. https://doi.org/10.4161/epi.23470

Choi, S. W., Mak, T. S.-H., & O'Reilly, P. F. (2020). Tutorial: A guide to performing polygenic risk score analyses. *Nature Protocols*, *15*(9), 2759–2772. https://doi.org/10.1038/s41596-020-0353-1

Christensen, Z. P., Freedman, E. G., & Foxe, J. J. (2021). Caffeine exposure in utero is associated with structural brain alterations and deleterious neurocognitive outcomes in 9–10 year old children. *Neuropharmacology*, *186*, 108479. https://doi.org/10.1016/j.neuropharm.2021.108479

Ciuculete, D. M., Boström, A. E., Tuunainen, A.-K., Sohrabi, F., Kular, L., Jagodic, M., Voisin, S., Mwinyi, J., & Schiöth, H. B. (2018). Changes in methylation within the STK32B promoter are associated with an increased risk for generalized anxiety disorder in adolescents. *Journal of Psychiatric Research*, *102*, 44–51. https://doi.org/10.1016/j.jpsychires.2018.03.008

Clark, S. L., Hattab, M. W., Chan, R. F., Shabalin, A. A., Han, L. K. M., Zhao, M., Smit, J. H., Jansen, R., Milaneschi, Y., Xie, L. Y., van Grootheest, G., Penninx, B. W. J. H., Aberg, K. A., & van den Oord, E. J. C. G. (2020). A methylation study of long-term depression risk. *Molecular Psychiatry*, *25*(6), 1334–1343. https://doi.org/10.1038/s41380-019-0516-z

Coleman, T., Chamberlain, C., Davey, M.-A., Cooper, S. E., & Leonardi-Bee, J. (2015). Pharmacological interventions for promoting smoking cessation during pregnancy. *Cochrane Database of Systematic Reviews*, *12*. https://doi.org/10.1002/14651858.CD010078.pub2

Coleman, T., Cooper, S., Thornton, J. G., Grainge, M. J., Watts, K., Britton, J., & Lewis, S. (2012). A Randomized Trial of Nicotine-Replacement Therapy Patches in Pregnancy. *New England Journal of Medicine*, *366*(9), 808–818. https://doi.org/10.1056/NEJMoa1109582

Conneely, K. N., & Boehnke, M. (2007). So Many Correlated Tests, So Little Time! Rapid Adjustment of P Values for Multiple Correlated Tests. *The American Journal of Human Genetics*, *81*(6), 1158–1168. https://doi.org/10.1086/522036

Cooper, D. L., Petherick, E. S., & Wright, J. (2013). The association between binge drinking and birth outcomes: Results from the Born in Bradford

cohort study. *J Epidemiol Community Health*, *67*(10), 821–828. https://doi.org/10.1136/jech-2012-202303

Cornelis, M., Kacprowski, T., Menni, C., Gustafsson, S., Pivin, E., Adamski, J., Artati, A., Eap, C. B., Ehret, G., Friedrich, N., Ganna, A., Guessous, I., Homuth, G., Lind, L., Magnusson, P. K., Mangino, M., Pedersen, N. L., Pietzner, M., Suhre, K., … Ingelsson, E. (2016). Genome-wide association study of caffeine metabolites provides new insights to caffeine metabolism and dietary caffeine-consumption behavior. *Human Molecular Genetics*, ddw334. https://doi.org/10.1093/hmg/ddw334

Cornelis, M., & Munafo, M. (2018). Mendelian Randomization Studies of Coffee and Caffeine Consumption. *Nutrients*, *10*(10), 1343. https://doi.org/10.3390/nu10101343

Cortessis, V. K., Thomas, D. C., Levine, A. J., Breton, C. V., Mack, T. M., Siegmund, K. D., Haile, R. W., & Laird, P. W. (2012). Environmental epigenetics: Prospects for studying epigenetic mediation of exposure–response relationships. *Human Genetics*, *131*(10), 1565–1589. https://doi.org/10.1007/s00439-012-1189-8

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry*, *150*(6), 782–786.

Crean, A. J., & Bonduriansky, R. (2014). What is a paternal effect? *Trends in Ecology & Evolution*, *29*(10), 554–559. https://doi.org/10.1016/j.tree.2014.07.009

Creswell, C., Waite, P., & Cooper, P. J. (2014). Assessment and management of anxiety disorders in children and adolescents. *Archives of Disease in Childhood*, *99*(7), 674–678. https://doi.org/10.1136/archdischild-2013-303768

Crone, M. R., & Reijneveld, S. A. (2007). The association of behavioural and emotional problems with tobacco use in adolescence. *Addictive Behaviors*, *32*(8), 1692–1698. https://doi.org/10.1016/j.addbeh.2006.11.006

Cummings, C. M., Caporino, N. E., & Kendall, P. C. (2014). Comorbidity of Anxiety and Depression in Children and Adolescents: 20 Years After. *Psychological Bulletin*, *140*(3), 816–845. https://doi.org/10.1037/a0034733

Daly, M., & Egan, M. (2017). Childhood cognitive ability and smoking initiation, relapse and cessation throughout adulthood: Evidence from two British cohort studies. *Addiction*, *112*(4), 651–659. https://doi.org/10.1111/add.13554

Davey Smith, G., & Ebrahim, S. (2003). 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease?*. *International Journal of Epidemiology*, *32*(1), 1–22. https://doi.org/10.1093/ije/dyg070

Davey Smith, G., & Hemani, G. (2014). Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics*, *23*(R1), R89–R98. https://doi.org/10.1093/hmg/ddu328

Davey Smith, G., Lawlor, D. A., Harbord, R., Timpson, N., Day, I., & Ebrahim, S. (2007). Clustered environments and randomized genes: A fundamental distinction between conventional and genetic epidemiology. *PLoS Med*, *4*(12), e352.

Davey-Smith, G. (2008). Assessing intrauterine influences on offspring health outcomes: Can epidemiological studies yield robust findings? *Basic & Clinical Pharmacology & Toxicology*, *102*(2), 245–256. WorldCat.org. https://doi.org/10.1111/j.1742-7843.2007.00191.x

Davies, M. N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., Coarfa, C., Harris, R. A., Milosavljevic, A., Troakes, C., Al-Sarraj, S., Dobson, R., Schalkwyk, L. C., & Mill, J. (2012). Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biology*, *13*(6), R43. https://doi.org/10.1186/gb-2012-13-6-r43

Davies, N. M., Holmes, M. V., & Davey Smith, G. (2018). Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ*, k601. https://doi.org/10.1136/bmj.k601

Delpisheh, A., Kelly, Y., Rizwan, S., Attia, E., Drammond, S., & Brabin, B. J. (2007). Population attributable risk for adverse pregnancy outcomes related to smoking in adolescents and adults. *Public Health*, *121*(11), 861–868. https://doi.org/10.1016/j.puhe.2007.03.015

Del-Ponte, B., Santos, I. S., Tovo-Rodrigues, L., Anselmi, L., Munhoz, T. N., & Matijasevich, A. (2016). Caffeine consumption during pregnancy and

ADHD at the age of 11 years: A birth cohort study. *BMJ Open*, *6*(12). https://doi.org/10.1136/bmjopen-2016-012749

Dempster, E. L., Wong, C. C. Y., Lester, K. J., Burrage, J., Gregory, A. M., Mill, J., & Eley, T. C. (2014). Genome-wide Methylomic Analysis of Monozygotic Twins Discordant for Adolescent Depression. *Biological Psychiatry*, *76*(12), 977–983. https://doi.org/10.1016/j.biopsych.2014.04.013

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, *13*(3), 595–605. https://doi.org/10.1017/S0033291700048017

Derrogatis, L., Lipman, R., & Covi, I. (1973). The SCL-90: An outpatient psychiatric rating scale. *Psychopharmacology Bulletin*, *9*(1), 13–28.

Diamanti, A., Papadakis, S., Schoretsaniti, S., Rovina, N., Vivilaki, V., Gratziou, C., & Katsaounou, P. A. (2019). Smoking cessation in pregnancy: An update for maternity care practitioners. *Tobacco Induced Diseases*, *17*. https://doi.org/10.18332/tid/109906

Dick, B., & Ferguson, B. J. (2015). Health for the World's Adolescents: A Second Chance in the Second Decade. *Journal of Adolescent Health*, *56*(1), 3–6. https://doi.org/10.1016/j.jadohealth.2014.10.260

DiClemente, C., Dolan-Mullen, P., & Windsor, R. (2000). The process of pregnancy smoking cessation: Implications for interventions. *Tobacco Control*, *9*(Suppl 3), iii16–iii21. https://doi.org/10.1136/tc.9.suppl_3.iii16

Diemer, E. W., Labrecque, J. A., Neumann, A., Tiemeier, H., & Swanson, S. A. (2020). Mendelian randomisation approaches to the study of prenatal exposures: A systematic review. *Paediatric and Perinatal Epidemiology*.

Dolan, C. V., Geels, L., Vink, J. M., van Beijsterveldt, C. E. M., Neale, M. C., Bartels, M., & Boomsma, D. I. (2016). Testing Causal Effects of Maternal Smoking During Pregnancy on Offspring's Externalizing and Internalizing Behavior. *Behavior Genetics*, *46*(3), 378–388. https://doi.org/10.1007/s10519-015-9738-2

D'Onofrio, B. M., Van Hulle, C. A., Waldman, I. D., Rodgers, J. L., Harden, K. P., Rathouz, P. J., & Lahey, B. B. (2008). Smoking during pregnancy and offspring externalizing problems: An exploration of genetic and environmental confounds. *Development and Psychopathology*, *20*(1), 139–

164. https://doi.org/10.1017/S0954579408000072

Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, *11*(1), 587. https://doi.org/10.1186/1471-2105-11-587

Dudbridge, F. (2016). Polygenic Epidemiology. *Genetic Epidemiology*, *40*(4), 268–272. https://doi.org/10.1002/gepi.21966

Duko, B., Ayano, G., Pereira, G., Betts, K., & Alati, R. (2020). Prenatal tobacco use and the risk of mood disorders in offspring: A systematic review and meta-analysis. *Social Psychiatry and Psychiatric Epidemiology*, *55*(12), 1549–1562. https://doi.org/10.1007/s00127-020-01949-y

Duncan, L. E., Ostacher, M., & Ballon, J. (2019). How genome-wide association studies (GWAS) made traditional candidate gene studies obsolete. *Neuropsychopharmacology*, *44*(9), 1518–1523. https://doi.org/10.1038/s41386-019-0389-5

Durbin, R. M., Altshuler, D., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Collins, F. S., De La Vega, F. M., Donnelly, P., Egholm, M., Flicek, P., Gabriel, S. B., Gibbs, R. A., Knoppers, B. M., Lander, E. S., Lehrach, H., Mardis, E. R., McVean, G. A., … The Translational Genomics Research Institute. (2010). A map of human genome variation from population-scale sequencing. *Nature*, *467*(7319), 1061–1073. https://doi.org/10.1038/nature09534

Duthie, L., & Reynolds, R. M. (2013). Changes in the Maternal Hypothalamic-Pituitary-Adrenal Axis in Pregnancy and Postpartum: Influences on Maternal and Fetal Outcomes. *Neuroendocrinology*, *98*(2), 106–115. https://doi.org/10.1159/000354702

Easey, K. E., & Sharp, G. C. (2021). The impact of paternal alcohol, tobacco, caffeine use and physical activity on offspring mental health: A systematic review and meta-analysis. *MedRxiv*, 2021.03.02.21252760. https://doi.org/10.1101/2021.03.02.21252760

Easey, K. E., Wootton, R. E., Sallis, H. M., Haan, E., Schellhas, L., Munafò, M. R., Timpson, N. J., & Zuccolo, L. (2021). Characterization of alcohol polygenic risk scores in the context of mental health outcomes: Within-individual and intergenerational analyses in the Avon Longitudinal Study

of Parents and Children. *Drug and Alcohol Dependence*, *221*, 108654. https://doi.org/10.1016/j.drugalcdep.2021.108654

Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V. K., Attwood, J., Burger, M., Burton, J., Cox, T. V., Davies, R., Down, T. A., Haefliger, C., Horton, R., Howe, K., Jackson, D. K., Kunde, J., Koenig, C., Liddle, J., Niblett, D., Otto, T., … Beck, S. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature Genetics*, *38*(12), 1378–1385. https://doi.org/10.1038/ng1909

Edgar, R. D., Jones, M. J., Meaney, M. J., Turecki, G., & Kobor, M. S. (2017). BECon: A tool for interpreting DNA methylation findings from blood in the context of brain. *Translational Psychiatry*, *7*(8), e1187–e1187. https://doi.org/10.1038/tp.2017.171

EFSA Panel on Dietetic Products, Nutrition, & Allergies (NDA). (2015). Scientific Opinion on the safety of caffeine. *EFSA Journal*, *13*(5), 4102.

Ehret, G. B. (2010). Genome-Wide Association Studies: Contribution of Genomics to Understanding Blood Pressure and Essential Hypertension. *Current Hypertension Reports*, *12*(1), 17–25. https://doi.org/10.1007/s11906-009-0086-6

Eilertsen, E. M., Hannigan, L. J., McAdams, T. A., Rijsdijk, F. V., Czajkowski, N., Reichborn-Kjennerud, T., Ystrom, E., & Gjerde, L. C. (2020). Parental Prenatal Symptoms of Depression and Offspring Symptoms of ADHD: A Genetically Informed Intergenerational Study. *Journal of Attention Disorders*, 1087054720914386. https://doi.org/10.1177/1087054720914386

Ekblad, M., Gissler, M., Lehtonen, L., & Korkeila, J. (2010). Prenatal Smoking Exposure and the Risk of Psychiatric Morbidity Into Young Adulthood. *Archives of General Psychiatry*, *67*(8), 841–849. https://doi.org/10.1001/archgenpsychiatry.2010.92

Eley, T. C. (1999). Behavioral genetics as a tool for developmental psychology: Anxiety and depression in children and adolescents. *Clinical Child and Family Psychology Review*, *2*(1), 21–36.

Elkins, I. J., McGue, M., & Iacono, W. G. (2007). Prospective effects of attention-deficit/hyperactivity disorder, conduct disorder, and sex on adolescent substance use and abuse. *Archives of General Psychiatry*, *64*(10), 1145–

1152. https://doi.org/10.1001/archpsyc.64.10.1145

Emeny, R. T., Baumert, J., Zannas, A. S., Kunze, S., Wahl, S., Iurato, S., Arloth, J., Erhardt, A., Balsevich, G., Schmidt, M. V., Weber, P., Kretschmer, A., Pfeiffer, L., Kruse, J., Strauch, K., Roden, M., Herder, C., Koenig, W., Gieger, C., … Ladwig, K.-H. (2018). Anxiety Associated Increased CpG Methylation in the Promoter of Asb1: A Translational Approach Evidenced by Epidemiological and Clinical Studies and a Murine Model. *Neuropsychopharmacology*, *43*(2), 342–353. https://doi.org/10.1038/npp.2017.102

England, L., Grauman, A., Qian, C., Wilkins, D., Schisterman, E., Yu, K., & Levine, R. (2007). Misclassification of maternal smoking status and its effects on an epidemiologic study of pregnancy outcomes. *Nicotine & Tobacco Research*, *9*(10), 1005–1013. https://doi.org/10.1080/14622200701491255

Ershoff, D. H. (2000). Predictors of intentions to stop smoking early in prenatal care. *Tobacco Control*, *9*(90003), 41iii–45. https://doi.org/10.1136/tc.9.suppl_3.iii41

Erskine, H. E., Moffitt, T. E., Copeland, W. E., Costello, E. J., Ferrari, A. J., Patton, G., Degenhardt, L., Vos, T., Whiteford, H. A., & Scott, J. G. (2015). A heavy burden on young minds: The global burden of mental and substance use disorders in children and youth. *Psychological Medicine*, *45*(7), 1551–1563. https://doi.org/10.1017/S0033291714002888

Evans, D. M., Moen, G.-H., Hwang, L.-D., Lawlor, D. A., & Warrington, N. M. (2019). Elucidating the role of maternal environmental exposures on offspring health and disease using two-sample Mendelian randomization. *International Journal of Epidemiology*, *48*(3), 861–875. https://doi.org/10.1093/ije/dyz019

Fagny, M., Patin, E., MacIsaac, J. L., Rotival, M., Flutre, T., Jones, M. J., Siddle, K. J., Quach, H., Harmant, C., McEwen, L. M., Froment, A., Heyer, E., Gessain, A., Betsem, E., Mouguiama-Daouda, P., Hombert, J.-M., Perry, G. H., Barreiro, L. B., Kobor, M. S., & Quintana-Murci, L. (2015). The epigenomic landscape of African rainforest hunter-gatherers and farmers. *Nature Communications*, *6*(1), 10047. https://doi.org/10.1038/ncomms10047

Fang, X., Mei, W., Barbazuk, W. B., Rivkees, S. A., & Wendler, C. C. (2014). Caffeine exposure alters cardiac gene expression in embryonic cardiomyocytes. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *307*(12), R1471–R1487. https://doi.org/10.1152/ajpregu.00307.2014

Farrow, A., Shea, K. M., & Little, R. E. (1998). Birthweight of term infants and maternal occupation in a prospective cohort of pregnant women. The ALSPAC Study Team. *Occupational and Environmental Medicine*, *55*(1), 18–23.

Fedorov, V., Mannino, F., & Zhang, R. (2009). Consequences of dichotomization. *Pharmaceutical Statistics*, *8*(1), 50–61. https://doi.org/10.1002/pst.331

Felix, J. F., Joubert, B. R., Baccarelli, A. A., Sharp, G. C., Almqvist, C., Annesi-Maesano, I., Arshad, H., Baïz, N., Bakermans-Kranenburg, M. J., Bakulski, K. M., Binder, E. B., Bouchard, L., Breton, C. V., Brunekreef, B., Brunst, K. J., Burchard, E. G., Bustamante, M., Chatzi, L., Cheng Munthe-Kaas, M., … London, S. J. (2018). Cohort Profile: Pregnancy And Childhood Epigenetics (PACE) Consortium. *International Journal of Epidemiology*, *47*(1), 22–23u. https://doi.org/10.1093/ije/dyx190

Ferrar, J., Griffith, G., Skirrow, C., Cashdollar, N., Taptiklis, N., Dobson, J., Cree, F., Cormack, F. K., Barnett, J. H., & Munafò, M. R. (2020). *Validating digital assessments for remote clinical research*. https://www.researchgate.net/profile/Jennifer-Ferrar/publication/347552688_Validating_digital_assessments_for_remote_clinical_research_Preprint/links/5ffd710092851c13fe06cadb/Validating-digital-assessments-for-remote-clinical-research-Preprint.pdf

Folkersen, L., Pain, O., Ingason, A., Werge, T., Lewis, C. M., & Austin, J. (2020). Impute.me: An Open-Source, Non-profit Tool for Using Data From Direct-to-Consumer Genetic Testing to Calculate and Interpret Polygenic Risk Scores. *Frontiers in Genetics*, *11*, 578. https://doi.org/10.3389/fgene.2020.00578

Fortin, J.-P., Labbe, A., Lemire, M., Zanke, B. W., Hudson, T. J., Fertig, E. J., Greenwood, C. M., & Hansen, K. D. (2014). Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biology*, *15*(11), 503. https://doi.org/10.1186/s13059-014-0503-2

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y.-Z., … Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(30), 10604–10609. https://doi.org/10.1073/pnas.0500398102

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., Macleod, J., Molloy, L., Ness, A., Ring, S., Nelson, S. M., & Lawlor, D. A. (2013). Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology*, *42*(1), 97–110. https://doi.org/10.1093/ije/dys066

Fredholm, B. B., Bättig, K., Holmén, J., Nehlig, A., & Zvartau, E. E. (1999). Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacological Reviews*, *51*(1), 83–133.

Fuchikami, M., Morinobu, S., Segawa, M., Okamoto, Y., Yamawaki, S., Ozaki, N., Inoue, T., Kusumi, I., Koyama, T., Tsuchiyama, K., & Terao, T. (2011). DNA Methylation Profiles of the Brain-Derived Neurotrophic Factor (BDNF) Gene as a Potent Diagnostic Biomarker in Major Depression. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0023881

Fuzik, J., Rehman, S., Girach, F., Miklosi, A. G., Korchynska, S., Arque, G., Romanov, R. A., Hanics, J., Wagner, L., Meletis, K., Yanagawa, Y., Kovacs, G. G., Alpár, A., Hökfelt, T. G. M., & Harkany, T. (2019). Brain-wide genetic mapping identifies the indusium griseum as a prenatal target of pharmacologically unrelated psychostimulants. *Proceedings of the National Academy of Sciences*, *116*(51), 25958–25967. https://doi.org/10.1073/pnas.1904006116

Gage, S. H., Bowden, J., Davey Smith, G., & Munafò, M. R. (2018). Investigating causality in associations between education and smoking: A two-sample Mendelian randomization study. *International Journal of Epidemiology*, *47*(4), 1131–1140. https://doi.org/10.1093/ije/dyy131

Gage, S. H., Munafò, M. R., & Davey Smith, G. (2016). Causal Inference in Developmental Origins of Health and Disease (DOHaD) Research. *Annual Review of Psychology*, *67*(1), 567–585. https://doi.org/10.1146/annurev-psych-122414-033352

Galanter, J. M., Gignoux, C. R., Oh, S. S., Torgerson, D., Pino-Yanes, M., Thakur, N., Eng, C., Hu, D., Huntsman, S., Farber, H. J., Avila, P. C., Brigino-Buenaventura, E., LeNoir, M. A., Meade, K., Serebrisky, D., Rodríguez-Cintrón, W., Kumar, R., Rodríguez-Santana, J. R., Seibold, M. A., … Zaitlen, N. (2017). Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *ELife*, *6*, e20532. https://doi.org/10.7554/eLife.20532

Garber, J., & Weersing, V. R. (2010). Comorbidity of Anxiety and Depression in Youth: Implications for Treatment and Prevention. *Clinical Psychology: Science and Practice*, *17*(4), 293–306. https://doi.org/10.1111/j.1468-2850.2010.01221.x

García-Giménez, J. L., Seco-Cervera, M., Tollefsbol, T. O., Romá-Mateo, C., Peiró-Chova, L., Lapunzina, P., & Pallardó, F. V. (2017). Epigenetic biomarkers: Current strategies and future challenges for their use in the clinical laboratory. *Critical Reviews in Clinical Laboratory Sciences*, *54*(7–8), 529–550. https://doi.org/10.1080/10408363.2017.1410520

Garner, J. P. (2014). The Significance of Meaning: Why Do Over 90% of Behavioral Neuroscience Results Fail to Translate to Humans, and What Can We Do to Fix It? *ILAR Journal*, *55*(3), 438–456. https://doi.org/10.1093/ilar/ilu047

Gartstein, M. A., Bridgett, D. J., Dishion, T. J., & Kaufman, N. K. (2009). Depressed mood and maternal report of child behavior problems: Another look at the depression–distortion hypothesis. *Journal of Applied Developmental Psychology*, *30*(2), 149–160. https://doi.org/10.1016/j.appdev.2008.12.001

Gaunt, T. R., Shihab, H. A., Hemani, G., Min, J. L., Woodward, G., Lyttleton, O., Zheng, J., Duggirala, A., McArdle, W. L., Ho, K., Ring, S. M., Evans, D. M., Davey Smith, G., & Relton, C. L. (2016). Systematic identification of genetic influences on methylation across the human life course. *Genome Biology*, *17*(1), 61. https://doi.org/10.1186/s13059-016-0926-z

Geeleher, P., Hartnett, L., Egan, L. J., Golden, A., Raja Ali, R. A., & Seoighe, C. (2013). Gene-set analysis is severely biased when applied to genome-wide methylation data. *Bioinformatics*, *29*(15), 1851–1857. https://doi.org/10.1093/bioinformatics/btt311

Gervin, K., Page, C. M., Aass, H. C. D., Jansen, M. A., Fjeldstad, H. E., Andreassen, B. K., Duijts, L., van Meurs, J. B., van Zelm, M. C., Jaddoe, V. W., Nordeng, H., Knudsen, G. P., Magnus, P., Nystad, W., Staff, A. C., Felix, J. F., & Lyle, R. (2016). Cell type specific DNA methylation in cord blood: A 450K-reference data set and cell count-based validation of estimated cell type composition. *Epigenetics*, *11*(9), 690–698. https://doi.org/10.1080/15592294.2016.1214782

Giglia, R. C., Binns, C. W., & Alfonso, H. S. (2006). Which women stop smoking during pregnancy and the effect on breastfeeding duration. *BMC Public Health*, *6*(1), 195. https://doi.org/10.1186/1471-2458-6-195

Gignon, M., Ganry, O., Jardé, O., & Manaouil, C. (2013). The Precautionary Principle: Is it Safe. *European Journal of Health Law*, *20*(3), 261–270. https://doi.org/10.1163/15718093-12341272

Gillespie, S. L., Mitchell, A. M., Kowalsky, J. M., & Christian, L. M. (2018). Maternal parity and perinatal cortisol adaptation: The role of pregnancy-specific distress and implications for postpartum mood. *Psychoneuroendocrinology*, *97*, 86–93. https://doi.org/10.1016/j.psyneuen.2018.07.008

Goddard, A. W., Ball, S. G., Martinez, J., Robinson, M. J., Yang, C. R., Russell, J. M., & Shekhar, A. (2010). Current perspectives of the roles of the central norepinephrine system in anxiety and depression. *Depression and Anxiety*, *27*(4), 339–350. https://doi.org/10.1002/da.20642

Goedhart, G., van der Wal, M. F., Cuijpers, P., & Bonsel, G. J. (2009). Psychosocial problems and continued smoking during pregnancy. *Addictive Behaviors*, *34*(4), 403–406. https://doi.org/10.1016/j.addbeh.2008.11.006

Goldman, M. (2008). Why is multiple testing a problem. *Statistics for Bioinformatics*.

Goldsack, J. C., Izmailova, E. S., Menetski, J. P., Hoffmann, S. C., Groenen, P. M. A., & Wagner, J. A. (2020). Remote digital monitoring in clinical trials

in the time of COVID-19. *Nature Reviews Drug Discovery*, *19*(6), 378–379. https://doi.org/10.1038/d41573-020-00094-0

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Goodman, R., & Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is small beautiful? *Journal of Abnormal Child Psychology*, *27*(1), 17–24.

Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well-Being Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *41*(5), 645–655.

Goodman, S. H., Rouse, M. H., Connell, A. M., Broth, M. R., Hall, C. M., & Heyward, D. (2011). Maternal Depression and Child Psychopathology: A Meta-Analytic Review. *Clinical Child and Family Psychology Review*, *14*(1), 1–27. https://doi.org/10.1007/s10567-010-0080-1

Goodwin, R. D., Cheslack-Postava, K., Nelson, D. B., Smith, P. H., Hasin, D. S., Janevic, T., Bakoyiannis, N., & Wall, M. M. (2017). Serious Psychological Distress and Smoking During Pregnancy in the United States: 2008–2014. *Nicotine & Tobacco Research*, *19*(5), 605–614. https://doi.org/10.1093/ntr/ntw323

Gottschalk, M. G., & Domschke, K. (2017). Genetics of generalized anxiety disorder and related traits. *Dialogues in Clinical Neuroscience*, *19*(2), 159–168.

Green, H., Maginnity, A., Meltzer, H., Ford, T., Goodman, R. (2005). *Mental health of children and young people in Great Britain*. The Stationary Office. https://files.digital.nhs.uk/publicationimport/pub06xxx/pub06116/ment-heal-chil-youn-peop-gb-2004-rep1.pdf

Gresham, E., Byles, J. E., Bisquera, A., & Hure, A. J. (2014). Effects of dietary interventions on neonatal and infant outcomes: A systematic review and meta-analysis. *The American Journal of Clinical Nutrition*, *100*(5), 1298–1321. https://doi.org/10.3945/ajcn.113.080655

Griffith, G. J., Morris, T. T., Tudball, M. J., Herbert, A., Mancano, G., Pike, L.,

Sharp, G. C., Sterne, J., Palmer, T. M., Davey Smith, G., Tilling, K., Zuccolo, L., Davies, N. M., & Hemani, G. (2020). Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nature Communications*, *11*. https://doi.org/10.1038/s41467-020-19478-2

Grosso, L. M., & Bracken, M. B. (2005). Caffeine Metabolism, Genetics, and Perinatal Outcomes: A Review of Exposure Assessment Considerations during Pregnancy. *Annals of Epidemiology*, *15*(6), 460–466. https://doi.org/10.1016/j.annepidem.2004.12.011

Gurdasani, D., Barroso, I., Zeggini, E., & Sandhu, M. S. (2019). Genomics of disease risk in globally diverse populations. *Nature Reviews Genetics*, *20*(9), 520–535. https://doi.org/10.1038/s41576-019-0144-0

Gutman, L. M., & Codiroli McMaster, N. (2020). Gendered Pathways of Internalizing Problems from Early Childhood to Adolescence and Associated Adolescent Outcomes. *Journal of Abnormal Child Psychology*, *48*(5), 703–718. https://doi.org/10.1007/s10802-020-00623-w

Gutteling, B. M., Weerth, C. de, & Buitelaar, J. K. (2005). Prenatal stress and children's cortisol reaction to the first day of school. *Psychoneuroendocrinology*, *30*(6), 541–549. https://doi.org/10.1016/j.psyneuen.2005.01.002

Guxens, M., Ballester, F., Espada, M., Fernández, M. F., Grimalt, J. O., Ibarluzea, J., Olea, N., Rebagliato, M., Tardón, A., Torrent, M., Vioque, J., Vrijheid, M., Sunyer, J., & INMA Project. (2012). Cohort Profile: The INMA--INfancia y Medio Ambiente--(Environment and Childhood) Project. *International Journal of Epidemiology*, *41*(4), 930–940. https://doi.org/10.1093/ije/dyr054

Hall, M. A., Verma, A., Brown-Gentry, K. D., Goodloe, R., Boston, J., Wilson, S., McClellan, B., Sutcliffe, C., Dilks, H. H., Gillani, N. B., Jin, H., Mayo, P., Allen, M., Schnetz-Boutaud, N., Crawford, D. C., Ritchie, M. D., & Pendergrass, S. A. (2014). Detection of Pleiotropy through a Phenome-Wide Association Study (PheWAS) of Epidemiologic Data as Part of the Environmental Architecture for Genes Linked to Environment (EAGLE) Study. *PLoS Genetics*, *10*(12). https://doi.org/10.1371/journal.pgen.1004678

Hamer, D., & Sirota, L. (2000). Beware the chopsticks gene. *Molecular*

Psychiatry, *5*(1), 11–13. https://doi.org/10.1038/sj.mp.4000662

Hammerton, G., & Munafò, M. R. (2021). Causal inference with observational data: The need for triangulation of evidence. *Psychological Medicine*, 1–16. https://doi.org/10.1017/S0033291720005127

Hannon, E., Lunnon, K., Schalkwyk, L., & Mill, J. (2015). Interindividual methylomic variation across blood, cortex, and cerebellum: Implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*, *10*(11), 1024–1032. https://doi.org/10.1080/15592294.2015.1100786

Harrison, R., Munafo, M. R., Smith, G. D., & Wootton, R. E. (2019). Examining the effect of smoking on suicidal ideation and attempts: A triangulation of epidemiological approaches. *MedRxiv*, 19007013. https://doi.org/10.1101/19007013

Heiss, J. A., & Just, A. C. (2019). Improved filtering of DNA methylation microarray data by detection p values and its impact on downstream analyses. *Clinical Epigenetics*, *11*. https://doi.org/10.1186/s13148-019-0615-3

Hellwege, J., Keaton, J., Giri, A., Gao, X., Velez Edwards, D. R., & Edwards, T. L. (2017). Population Stratification in Genetic Association Studies. *Current Protocols in Human Genetics*, *95*, 1.22.1-1.22.23. https://doi.org/10.1002/cphg.48

Hemani, G., Bowden, J., & Davey Smith, G. (2018). Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Human Molecular Genetics*, *27*(R2), R195–R208. https://doi.org/10.1093/hmg/ddy163

Hemani, G., Tilling, K., & Smith, G. D. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genetics*, *13*(11), e1007081. https://doi.org/10.1371/journal.pgen.1007081

Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., Tan, V. Y., Yarmolinsky, J., Shihab, H. A., Timpson, N. J., Evans, D. M., Relton, C., Martin, R. M., Davey Smith, G., Gaunt, T. R., & Haycock, P. C. (2018). The MR-Base platform supports systematic causal inference across the human phenome. *ELife*, *7*, e34408. https://doi.org/10.7554/eLife.34408

Hennigan, K. M., O’Keefe, M., Noether, C. D., Rinehart, D. J., & Russell, L. A.

(2006). Through a Mother's Eyes: Sources of Bias When Mothers with Co-occurring Disorders Assess Their Children. *The Journal of Behavioral Health Services & Research*, *33*(1), 87–104. https://doi.org/10.1007/s11414-005-9005-z

Heude, B., Forhan, A., Slama, R., Douhaud, L., Bedel, S., Saurel-Cubizolles, M.-J., Hankard, R., Thiebaugeorges, O., De Agostini, M., Annesi-Maesano, I., Kaminski, M., & Charles, M.-A. (2016). Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *International Journal of Epidemiology*, *45*(2), 353–363. https://doi.org/10.1093/ije/dyv151

Hicks, B. M., Clark, D. A., Deak, J. D., Liu, M., Durbin, C. E., Schaefer, J. D., Wilson, S., Iacono, W. G., McGue, M., & Vrieze, S. I. (2020). Polygenic Risk Score for Smoking is associated with Externalizing Psychopathology and Disinhibited Personality Traits but not Internalizing Psychopathology in Adolescence. *BioRxiv*, 2020.07.29.227405. https://doi.org/10.1101/2020.07.29.227405

Hillary, R. F., & Marioni, R. E. (2021). MethylDetectR: a software for methylation-based health profiling. *Wellcome Open Research*, *5*, 283. https://doi.org/10.12688/wellcomeopenres.16458.2

Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, *6*(2), 95–108. https://doi.org/10.1038/nrg1521

Hoek, H. W., Brown, A. S., & Susser, E. (1998). The Dutch Famine and schizophrenia spectrum disorders. *Social Psychiatry and Psychiatric Epidemiology*, *33*(8), 373–379. https://doi.org/10.1007/s001270050068

Hofman, A., Jaddoe, V. W. V., Mackenbach, J. P., Moll, H. A., Snijders, R. F. M., Steegers, E. A. P., Verhulst, F. C., Witteman, J. C. M., & Büller, H. A. (2004). Growth, development and health from early fetal life until young adulthood: The Generation R Study. *Paediatric and Perinatal Epidemiology*, *18*(1), 61–72. https://doi.org/10.1111/j.1365-3016.2003.00521.x

Hompes, T., Izzi, B., Gellens, E., Morreels, M., Fieuws, S., Pexsters, A., Schops, G., Dom, M., Van Bree, R., Freson, K., Verhaeghe, J., Spitz, B., Demyttenaere, K., Glover, V., Van den Bergh, B., Allegaert, K., & Claes,

S. (2013). Investigating the influence of maternal cortisol and emotional state during pregnancy on the DNA methylation status of the glucocorticoid receptor gene (NR3C1) promoter region in cord blood. *Journal of Psychiatric Research*, *47*(7), 880–891. https://doi.org/10.1016/j.jpsychires.2013.03.009

Höök, B., Cederblad, M., & Berg, R. (2006). Prenatal and postnatal maternal smoking as risk factors for preschool children's mental health. *Acta Paediatrica*, *95*(6), 671–677. https://doi.org/10.1111/j.1651-2227.2006.tb02314.x

Horvath, S., Zhang, Y., Langfelder, P., Kahn, R. S., Boks, M. P., van Eijk, K., van den Berg, L. H., & Ophoff, R. A. (2012). Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biology*, *13*(10), R97. https://doi.org/10.1186/gb-2012-13-10-r97

Horwood, J., Thomas, K., Duffy, L., Gunnell, D., Hollis, C., Lewis, G., Thompson, A., Wolke, D., Zammitt, S., & Harrison, G. (2008). P0305 - Frequency of psychosis-like symptoms in a non-clinical population of 12 year olds: Results from the Alspac birth cohort. *European Psychiatry*, *23*(S2), S282–S282. https://doi.org/10.1016/j.eurpsy.2008.01.595

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., Wiencke, J. K., & Kelsey, K. T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, *13*(1). https://doi.org/10.1186/1471-2105-13-86

Houseman, E. A., Molitor, J., & Marsit, C. J. (2014). Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics*, *30*(10), 1431–1439. https://doi.org/10.1093/bioinformatics/btu029

Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., … McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, *22*(3), 343–352. https://doi.org/10.1038/s41593-018-0326-7

Howard, D. M., Adams, M. J., Shirali, M., Clarke, T.-K., Marioni, R. E., Davies,

G., Coleman, J. R. I., Alloza, C., Shen, X., Barbu, M. C., Wigmore, E. M., Gibson, J., Hagenaars, S. P., Lewis, C. M., Ward, J., Smith, D. J., Sullivan, P. F., Haley, C. S., Breen, G., … McIntosh, A. M. (2018). Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nature Communications*, *9*(1), 1470. https://doi.org/10.1038/s41467-018-03819-3

Howard, L. M., Bekele, D., Rowe, M., Demilew, J., Bewley, S., & Marteau, T. M. (2013). Smoking cessation in pregnant women with mental disorders: A cohort and nested qualitative study. *BJOG: An International Journal of Obstetrics & Gynaecology*, *120*(3), 362–370. https://doi.org/10.1111/1471-0528.12059

Hoyt, A. T., Browne, M., Richardson, S., Romitti, P., & Druschel, C. (2014). Maternal Caffeine Consumption and Small for Gestational Age Births: Results from a Population-Based Case–Control Study. *Maternal and Child Health Journal*, *18*(6), 1540–1551. https://doi.org/10.1007/s10995-013-1397-4

Hughes, R. A., Heron, J., Sterne, J. A. C., & Tilling, K. (2019). Accounting for missing data in statistical analyses: Multiple imputation is not always the answer. *International Journal of Epidemiology*, *48*(4), 1294–1304. https://doi.org/10.1093/ije/dyz032

Hüls, A., & Czamara, D. (2020). Methodological challenges in constructing DNA methylation risk scores. *Epigenetics*, *15*(1–2), 1–11. https://doi.org/10.1080/15592294.2019.1644879

Hvolgaard Mikkelsen, S., Obel, C., Olsen, J., Niclasen, J., & Bech, B. H. (2017). Maternal Caffeine Consumption during Pregnancy and Behavioral Disorders in 11-Year-Old Offspring: A Danish National Birth Cohort Study. *The Journal of Pediatrics*, *189*, 120-127.e1. https://doi.org/10.1016/j.jpeds.2017.06.051

Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., Tung, J. Y., Hinds, D. A., Perlis, R. H., & Winslow, A. R. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics*, *48*(9), 1031–1036. https://doi.org/10.1038/ng.3623

Hylkema, M. N., & Blacquiere, M. J. (2009). Intrauterine Effects of Maternal

Smoking on Sensitization, Asthma, and Chronic Obstructive Pulmonary Disease. *Proceedings of the American Thoracic Society*, *6*(8), 660–662. https://doi.org/10.1513/pats.200907-065DP

Iacono, W. G., Malone, S. M., & McGue, M. (2008). Behavioral disinhibition and the development of early-onset addiction: Common and specific influences. *Annu. Rev. Clin. Psychol.*, *4*, 325–348.

Illingworth, R. S., & Bird, A. P. (2009). CpG islands – 'A rough guide'. *FEBS Letters*, *583*(11), 1713–1720. https://doi.org/10.1016/j.febslet.2009.04.012

Indredavik, M. S., Brubakk, A.-M., Romundstad, P., & Vik, T. (2007). Prenatal smoking exposure and psychiatric symptoms in adolescence. *Acta Paediatrica*, *96*(3), 377–382. https://doi.org/10.1111/j.1651-2227.2006.00148.x

Ismaili M'hamdi, H., de Beaufort, I., Jack, B., & Steegers, E. A. P. (2018). Responsibility in the age of Developmental Origins of Health and Disease (DOHaD) and epigenetics. *Journal of Developmental Origins of Health and Disease*, *9*(1), 58–62. https://doi.org/10.1017/S2040174417000654

Jacobson, N., & Newman, M. (2017). Anxiety and Depression as Bidirectional Risk Factors for One Another: A Meta-Analysis of Longitudinal Studies. *Psychological Bulletin*, *143*, 1155–1200. https://doi.org/10.1037/bul0000111

Jaffee, S. R., & Price, T. S. (2008). Genotype–environment correlations: Implications for determining the relationship between environmental exposures and psychiatric illness. *Psychiatry*, *7*(12), 496–499. https://doi.org/10.1016/j.mppsy.2008.10.002

James, J. E. (2020). Maternal caffeine consumption and pregnancy outcomes: A narrative review with implications for advice to mothers and mothers-to-be. *Maternal and Child Health*, *0*(0), 10.

Jami, E. S., Hammerschlag, A. R., Ip, H. F., Allegrini, A. G., Benyamin, B., Border, R., Diemer, E. W., Jiang, C., Karhunen, V., Lu, Y., Lu, Q., Mallard, T. T., Mishra, P. P., Nolte, I. M., Palviainen, T., Peterson, R. E., Sallis, H. M., Tate, A. E., Thiering, E., … Middeldorp, C. M. (2020). Genome-wide association meta-analysis of childhood and adolescent internalising symptoms. *MedRxiv*, 2020.09.11.20175026. https://doi.org/10.1101/2020.09.11.20175026

Jamieson, E., Korologou-Linden, R., Wootton, R. E., Guyatt, A. L., Battram, T., Burrows, K., ... & Richmond, R. C. (2020). Smoking, DNA methylation, and lung function: a mendelian randomization analysis to investigate causal pathways. *The American Journal of Human Genetics*, *106*(3), 315-326.

Janssens, A. C. J. W. (2019). Validity of polygenic risk scores: Are we measuring what we think we are? *Human Molecular Genetics*, *28*(R2), R143–R150. https://doi.org/10.1093/hmg/ddz205

Jauniaux, E., Gulbis, B., Acharya, G., Thiry, P., & Rodeck, C. (1999). Maternal tobacco exposure and cotinine levels in fetal fluids in the first half of pregnancy. *Obstetrics & Gynecology*, *93*(1), 25–29. https://doi.org/10.1016/S0029-7844(98)00318-4

Jones, M. J., Moore, S. R., & Kobor, M. S. (2018). Principles and challenges of applying epigenetic epidemiology to psychology. *Annual Review of Psychology*, *69*, 459–485.

Joubert, B. R., Felix, J. F., Yousefi, P., Bakulski, K. M., Just, A. C., Breton, C., Reese, S. E., Markunas, C. A., Richmond, R. C., Xu, C.-J., Küpers, L. K., Oh, S. S., Hoyo, C., Gruzieva, O., Söderhäll, C., Salas, L. A., Baïz, N., Zhang, H., Lepeule, J., … London, S. J. (2016). DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *American Journal of Human Genetics*, *98*(4), 680–696. https://doi.org/10.1016/j.ajhg.2016.02.019

Kamat, M. A., Blackshaw, J. A., Young, R., Surendran, P., Burgess, S., Danesh, J., Butterworth, A. S., & Staley, J. R. (2019). PhenoScanner V2: An expanded tool for searching human genotype–phenotype associations. *Bioinformatics*, *35*(22), 4851–4853. https://doi.org/10.1093/bioinformatics/btz469

Kaminsky, Z. A., Tang, T., Wang, S.-C., Ptak, C., Oh, G. H. T., Wong, A. H. C., Feldcamp, L. A., Virtanen, C., Halfvarson, J., Tysk, C., McRae, A. F., Visscher, P. M., Montgomery, G. W., Gottesman, I. I., Martin, N. G., & Petronis, A. (2009). DNA methylation profiles in monozygotic and dizygotic twins. *Nature Genetics*, *41*(2), 240–245. https://doi.org/10.1038/ng.286

Karabegović, I., Portilla-Fernandez, E., Li, Y., Ma, J., Maas, S. C. E., Sun, D.,

Hu, E. A., Kühnel, B., Zhang, Y., Ambatipudi, S., Fiorito, G., Huang, J., Castillo-Fernandez, J. E., Wiggins, K. L., de Klein, N., Grioni, S., Swenson, B. R., Polidoro, S., Treur, J. L., … Ghanbari, M. (2020). *Epigenome-wide association meta-analysis of DNA methylation with coffee and tea consumption* [Preprint]. Genomics. https://doi.org/10.1101/2020.04.15.042267

Keeler, J. F., & Robbins, T. W. (2011). Translating cognition from animals to humans. *Biochemical Pharmacology*, *81*(12), 1356–1366. https://doi.org/10.1016/j.bcp.2010.12.028

Kendler, K. S., Schmitt, E., Aggen, S. H., & Prescott, C. A. (2008). Genetic and Environmental Influences on Alcohol, Caffeine, Cannabis, and Nicotine Use From Early Adolescence to Middle Adulthood. *ARCH GEN PSYCHIATRY*, *65*(6), 9.

Kenney, M., & Müller, R. (2017). Of rats and women: Narratives of motherhood in environmental epigenetics. *BioSocieties*, *12*(1), 23–46. https://doi.org/10.1057/s41292-016-0002-7

Kentner, A. C., Cryan, J. F., & Brummelte, S. (2019). Resilience priming: Translational models for understanding resiliency and adaptation to early life adversity. *Developmental Psychobiology*, *61*(3), 350–375. https://doi.org/10.1002/dev.21775

Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B., & Wittchen, H.-U. (1998). The World Health Organization composite international diagnostic interview short-form (CIDI-SF). *International Journal of Methods in Psychiatric Research*, *7*(4), 171–185.

Khantzian, E. J. (1997). The self-medication hypothesis of substance use disorders: A reconsideration and recent applications. *Harvard Review of Psychiatry*, *4*(5), 231–244. https://doi.org/10.3109/10673229709030550

Khouja, J. N., Wootton, R. E., Taylor, A. E., Smith, G. D., & Munafò, M. R. (2021). Association of genetic liability to smoking initiation with e-cigarette use in young adults: A cohort study. *PLOS Medicine*, *18*(3), e1003555. https://doi.org/10.1371/journal.pmed.1003555

Kim-Cohen, J., Caspi, A., Moffitt, T. E., Harrington, H., Milne, B. J., & Poulton, R. (2003). Prior Juvenile Diagnoses in Adults With Mental Disorder: Developmental Follow-Back of a Prospective-Longitudinal Cohort.

Archives of General Psychiatry, *60*(7), 709. https://doi.org/10.1001/archpsyc.60.7.709

Klasen, D. H. (2000). Comparing the German Versions of the Strengths and Dif®culties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. *Adolescent Psychiatry*, *9*(4), 6.

Klebanoff, M. A., & Keim, S. A. (2015). Maternal Caffeine Intake During Pregnancy and Child Cognition and Behavior at 4 and 7 Years of Age. *American Journal of Epidemiology*, *182*(12), 1023–1032. https://doi.org/10.1093/aje/kwv136

Klengel, T., & Binder, E. B. (2015). Epigenetics of Stress-Related Psychiatric Disorders and Gene × Environment Interactions. *Neuron*, *86*(6), 1343–1357. https://doi.org/10.1016/j.neuron.2015.05.036

Knight, C., Knight, I., Mitchell, D., & Zepp, J. E. (2005). Beverage caffeine intake in US consumers and subpopulations of interest: Estimates from the Share of Intake Panel survey. *Food and Chemical Toxicology : An International Journal Published for the British Industrial Biological Research Association*, *42*, 1923–1930. https://doi.org/10.1016/j.fct.2004.05.002

Knopik, V. S., Marceau, K., Bidwell, L. C., & Rolan, E. (2018). Prenatal substance exposure and offspring development: Does DNA methylation play a role? *Neurotoxicology and Teratology*. https://doi.org/10.1016/j.ntt.2018.01.009

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., Benonisdottir, S., Oddsson, A., Halldorsson, B. V., Masson, G., Gudbjartsson, D. F., Helgason, A., Bjornsdottir, G., Thorsteinsdottir, U., & Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, *359*(6374), 424–428. https://doi.org/10.1126/science.aan6877

Koniak-Griffin, D., Logsdon, M. C., Hines-Martin, V., & Turner, C. C. (2006). Contemporary Mothering in a Diverse Society. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, *35*(5), 671–678. https://doi.org/10.1111/j.1552-6909.2006.00089.x

Kooijman, M. N., Kruithof, C. J., van Duijn, C. M., Duijts, L., Franco, O. H., van IJzendoorn, M. H., de Jongste, J. C., Klaver, C. C. W., van der Lugt, A.,

Mackenbach, J. P., Moll, H. A., Peeters, R. P., Raat, H., Rings, E. H. H. M., Rivadeneira, F., van der Schroeff, M. P., Steegers, E. A. P., Tiemeier, H., Uitterlinden, A. G., … Jaddoe, V. W. V. (2016). The Generation R Study: Design and cohort update 2017. *European Journal of Epidemiology*, *31*(12), 1243–1264. https://doi.org/10.1007/s10654-016-0224-9

Koolschijn, P. C. M. P., van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Crone, E. A. (2013). Hippocampal volume and internalizing behavior problems in adolescence. *European Neuropsychopharmacology*, *23*(7), 622–628. https://doi.org/10.1016/j.euroneuro.2012.07.001

Koskelainen, M., & Kaljonen, A. (2000). The Strengths and Dif®culties Questionnaire among Finnish school-aged children and adolescents. *Adolescent Psychiatry*, *9*(4), 8.

Kruithof, C. J., Kooijman, M. N., van Duijn, C. M., Franco, O. H., de Jongste, J. C., Klaver, C. C. W., Mackenbach, J. P., Moll, H. A., Raat, H., Rings, E. H. H. M., Rivadeneira, F., Steegers, E. A. P., Tiemeier, H., Uitterlinden, A. G., Verhulst, F. C., Wolvius, E. B., Hofman, A., & Jaddoe, V. W. V. (2014). The Generation R Study: Biobank update 2015. *European Journal of Epidemiology*, *29*(12), 911–927.

Kua, K. P., & Lee, S. W. H. (2017). Systematic review and meta-analysis of clinical outcomes of early caffeine therapy in preterm neonates. *British Journal of Clinical Pharmacology*, *83*(1), 180–191. https://doi.org/10.1111/bcp.13089

Kuan, P. F., & Chiang, D. Y. (2012). Integrating prior knowledge in multiple testing under dependence with applications to detecting differential DNA methylation. *Biometrics*, *68*(3), 774–783. https://doi.org/10.1111/j.1541-0420.2011.01730.x

Kumsta, R. (2019). The role of epigenetics for understanding mental health difficulties and its implications for psychotherapy research. *Psychology and Psychotherapy: Theory, Research and Practice*, *92*(2), 190–207. https://doi.org/10.1111/papt.12227

Kwok, M. K., Leung, G. M., & Schooling, C. M. (2016). Habitual coffee consumption and risk of type 2 diabetes, ischemic heart disease, depression and Alzheimer's disease: A Mendelian randomization study.

Scientific Reports, *6*(1). https://doi.org/10.1038/srep36500

Labonte, B., Yerko, V., Gross, J., Mechawar, N., Meaney, M. J., Szyf, M., & Turecki, G. (2012). Differential Glucocorticoid Receptor Exon 1B, 1C, and 1H Expression and Methylation in Suicide Completers with a History of Childhood Abuse. *Biological Psychiatry*, *72*(1), 41–48. https://doi.org/10.1016/j.biopsych.2012.01.034

Ladd-Acosta, C., & Fallin, M. D. (2015). The role of epigenetics in genetic and environmental epidemiology. *Epigenomics*, *8*(2), 271–283. https://doi.org/10.2217/epi.15.102

Lange, S., Probst, C., Gmel, G., Rehm, J., Burd, L., & Popova, S. (2017). Global Prevalence of Fetal Alcohol Spectrum Disorder Among Children and Youth. *JAMA Pediatrics*, *171*(10), 948–956. https://doi.org/10.1001/jamapediatrics.2017.1919

Lange, S., Probst, C., Rehm, J., & Popova, S. (2018). National, regional, and global prevalence of smoking during pregnancy in the general population: A systematic review and meta-analysis. *The Lancet Global Health*, *6*(7), e769–e776. https://doi.org/10.1016/S2214-109X(18)30223-7

Langley, K., Heron, J., Smith, G. D., & Thapar, A. (2012). Maternal and Paternal Smoking During Pregnancy and Risk of ADHD Symptoms in Offspring: Testing for Intrauterine Effects. *American Journal of Epidemiology*, *176*(3), 261–268. https://doi.org/10.1093/aje/kwr510

Lappalainen, T., & Greally, J. M. (2017). Associating cellular epigenetic models with human phenotypes. Nature Reviews Genetics, 18(7), 441–451. https://doi.org/10.1038/nrg.2017.32

Lara, D. R. (2010). Caffeine, Mental Health, and Psychiatric Disorders. *Journal of Alzheimer's Disease*, *20*(s1), S239–S248. https://doi.org/10.3233/JAD-2010-1378

Latimer, K., Wilson, P., Kemp, J., Thompson, L., Sim, F., Gillberg, C., Puckering, C., & Minnis, H. (2012). Disruptive behaviour disorders: A systematic review of environmental antenatal and early years risk factors. *Child: Care, Health and Development*, *38*(5), 611–628. https://doi.org/10.1111/j.1365-2214.2012.01366.x

Lavigne, J. V., Hopkins, J., Gouze, K. R., Bryant, F. B., LeBailly, S. A., Binns, H. J., & Lavigne, P. M. (2011). Is Smoking During Pregnancy a Risk Factor

for Psychopathology in Young Children? A Methodological Caveat and Report on Preschoolers. *Journal of Pediatric Psychology*, *36*(1), 10–24. https://doi.org/10.1093/jpepsy/jsq044

Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., & Smith, G. D. (2008). Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine*, *27*(8), 1133–1163. https://doi.org/10.1002/sim.3034

Lawlor, D. A., Lewcock, M., Rena-Jones, L., Rollings, C., Yip, V., Smith, D., Pearson, R. M., Johnson, L., Millard, L. A. C., Patel, N., Skinner, A., & Tilling, K. (2019). The second generation of The Avon Longitudinal Study of Parents and Children (ALSPAC-G2): A cohort profile. *Wellcome Open Research*, *4*. https://doi.org/10.12688/wellcomeopenres.15087.2

Lawlor, D. A., Richmond, R., Warrington, N., McMahon, G., Davey Smith, G., Bowden, J., & Evans, D. M. (2017). 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*, *2*, 11. https://doi.org/10.12688/wellcomeopenres.10567.1

Lawlor, D. A., Smith, G. D., Bruckdorfer, K. R., Kundu, D., & Ebrahim, S. (2004). Those confounded vitamins: What can we learn from the differences between observational versus randomised trial evidence? *The Lancet*, *363*(9422), 1724–1727. https://doi.org/10.1016/S0140-6736(04)16260-0

Lawlor, D. A., Tilling, K., & Davey Smith, G. (2016). Triangulation in aetiological epidemiology. *International Journal of Epidemiology*, *45*(6), 1866–1886. https://doi.org/10.1093/ije/dyw314

Lawlor, D. A., Wade, K., Borges, M. C., Palmer, T., Hartwig, F. P., Hemani, G., & Bowden, J. (2019). *A Mendelian Randomization dictionary: Useful definitions and descriptions for undertaking, understanding and interpreting Mendelian Randomization studies*.

Lawson, C. C., LeMasters, G. K., Levin, L. S., & Liu, J. H. (2002). Pregnancy hormone metabolite patterns, pregnancy symptoms, and coffee consumption. *American Journal of Epidemiology*, *156*(5), 428–437.

Lawson, C. C., LeMasters, G. K., & Wilson, K. A. (2004). Changes in caffeine

consumption as a signal of pregnancy. *Reproductive Toxicology*, *18*(5), 625–633. https://doi.org/10.1016/j.reprotox.2004.03.004

Lee, M. A., Cameron, O. G., & Greden, J. F. (1985). Anxiety and caffeine consumption in people with anxiety disorders. *Psychiatry Research*, *15*(3), 211–217. https://doi.org/10.1016/0165-1781(85)90078-2

Lee, S. H., Ripke, S., Neale, B. M., Faraone, S. V., Purcell, S. M., Perlis, R. H., Mowry, B. J., Thapar, A., Goddard, M. E., Witte, J. S., Absher, D., Agartz, I., Akil, H., Amin, F., Andreassen, O. A., Anjorin, A., Anney, R., Anttila, V., Arking, D. E., … Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics*, *45*(9), 984–994. https://doi.org/10.1038/ng.2711

Leek, J. T., Johnson, W. E., Parker, H. S., Fertig, E. J., Jaffe, A. E., Storey, J. D., Zhang, Y., & Torres, L. C. (2019). *sva: Surrogate Variable Analysis*.

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, *28*(6), 882–883. https://doi.org/10.1093/bioinformatics/bts034

Lewis, G., Pelosi, A. J., Araya, R., & Dunn, G. (1992). Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. *Psychological Medicine*, *22*(2), 465–486. https://doi.org/10.1017/s0033291700030415

Legendre, P., Legendre, L. (1998). *Numerical ecology* (2nd English ed.). Elsevier.

Lehne, B., Drong, A. W., Loh, M., Zhang, W., Scott, W. R., Tan, S.-T., Afzal, U., Scott, J., Jarvelin, M.-R., Elliott, P., McCarthy, M. I., Kooner, J. S., & Chambers, J. C. (2015). A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biology*, *16*(1), 37. https://doi.org/10.1186/s13059-015-0600-x

LeMoult, J., Humphreys, K. L., Tracy, A., Hoffmeister, J.-A., Ip, E., & Gotlib, I. H. (2020). Meta-analysis: Exposure to Early Life Stress and Risk for Depression in Childhood and Adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *59*(7), 842–855. https://doi.org/10.1016/j.jaac.2019.10.011

Lewis, S. J., Araya, R., Smith, G. D., Freathy, R., Gunnell, D., Palmer, T., & Munafò, M. (2011). Smoking Is Associated with, but Does Not Cause, Depressed Mood in Pregnancy – A Mendelian Randomization Study. *PLOS ONE*, *6*(7), e21689. https://doi.org/10.1371/journal.pone.0021689

Li, C., Gao, W., Gao, Y., Yu, C., Lv, J., Lv, R., Duan, J., Sun, Y., Guo, X., Cao, W., & Li, L. (2018). Age prediction of children and adolescents aged 6-17 years: An epigenome-wide analysis of DNA methylation. *Aging (Albany NY)*, *10*(5), 1015–1026. https://doi.org/10.18632/aging.101445

Li, M., D'Arcy, C., Li, X., Zhang, T., Joober, R., & Meng, X. (2019). What do DNA methylation studies tell us about depression? A systematic review. *Translational Psychiatry*, *9*. https://doi.org/10.1038/s41398-019-0412-y

Lieshout, R. J. V., & Krzeczkowski, J. E. (2016). Just DO(HaD) It! Testing the clinical potential of the DOHaD hypothesis to prevent mental disorders using experimental study designs. *Journal of Developmental Origins of Health and Disease*, *7*(6), 565–573. https://doi.org/10.1017/S2040174416000441

Linnet, K. M., Wisborg, K., Secher, N. J., Thomsen, P. H., Obel, C., Dalsgaard, S., & Henriksen, T. B. (2009). Coffee Consumption During Pregnancy and the Risk of Hyperkinetic Disorder and ADHD: A Prospective Cohort Study. *Obstetrical & Gynecological Survey*, *64*(4). https://doi.org/10.1097/01.ogx.0000345710.03145.80

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., Datta, G., Davila-Velderrain, J., McGuire, D., Tian, C., Zhan, X., Choquet, H., Docherty, A. R., Faul, J. D., Foerster, J. R., Fritsche, L. G., Gabrielsen, M. E., Gordon, S. D., Haessler, J., … Vrieze, S. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature Genetics*, *51*(2), 237–244. https://doi.org/10.1038/s41588-018-0307-5

Lokk, K., Modhukur, V., Rajashekar, B., Märtens, K., Mägi, R., Kolde, R., Koltšina, M., Nilsson, T. K., Vilo, J., Salumets, A., & Tõnisson, N. (2014). DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biology*, *15*(4), 3248. https://doi.org/10.1186/gb-2014-15-4-r54

Lynn, F. A., Alderdice, F. A., Crealey, G. E., & McElnay, J. C. (2011).

Associations between maternal characteristics and pregnancy-related stress among low-risk mothers: An observational cross-sectional study. *International Journal of Nursing Studies*, *48*(5), 620–627. https://doi.org/10.1016/j.ijnurstu.2010.10.002

Loomans, E. M., Hofland, L., van der Stelt, O., van der Wal, M. F., Koot, H. M., Van den Bergh, B. R. H., & Vrijkotte, T. G. M. (2012). Caffeine Intake During Pregnancy and Risk of Problem Behavior in 5- to 6-Year-Old Children. *PEDIATRICS*, *130*(2), e305–e313. https://doi.org/10.1542/peds.2011-3361

Lövkvist, C., Dodd, I. B., Sneppen, K., & Haerter, J. O. (2016). DNA methylation in human epigenomes depends on local topology of CpG sites. *Nucleic Acids Research*, *44*(11), 5123–5132. https://doi.org/10.1093/nar/gkw124

Ludwig, I. A., Clifford, M. N., Lean, M. E. J., Ashihara, H., & Crozier, A. (2014). Coffee: Biochemistry and potential impact on health. *Food Funct.*, *5*(8), 1695–1717. https://doi.org/10.1039/C4FO00042K

Lumey, L. H., Stein, A. D., Kahn, H. S., van der Pal-de Bruin, K. M., Blauw, G. J., Zybert, P. A., & Susser, E. S. (2007). Cohort Profile: The Dutch Hunger Winter Families Study. *International Journal of Epidemiology*, *36*(6), 1196–1204. https://doi.org/10.1093/ije/dym126

Lumley, J., Chamberlain, C., Dowswell, T., Oliver, S., Oakley, L., & Watson, L. (2009). Interventions for promoting smoking cessation during pregnancy. *The Cochrane Database of Systematic Reviews*, *3*, CD001055. https://doi.org/10.1002/14651858.CD001055.pub3

Lupien, S. J., McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience*, *10*(6), 434–445. https://doi.org/10.1038/nrn2639

Madley-Dowd, P., Rai, D., Zammit, S., & Heron, J. (2020). Simulations and directed acyclic graphs explained why assortative mating biases the prenatal negative control design. *Journal of Clinical Epidemiology*, *118*, 9–17. https://doi.org/10.1016/j.jclinepi.2019.10.008

Magnus, P., Birke, C., Vejrup, K., Haugan, A., Alsaker, E., Daltveit, A. K., Handal, M., Haugen, M., Høiseth, G., Knudsen, G. P., Paltiel, L., Schreuder, P., Tambs, K., Vold, L., & Stoltenberg, C. (2016). Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa).

International Journal of Epidemiology, *45*(2), 382–388. https://doi.org/10.1093/ije/dyw029

Magnus, P., Irgens, L. M., Haug, K., Nystad, W., Skjærven, R., & Stoltenberg, C. (2006). Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiology*, *35*(5), 1146–1150. https://doi.org/10.1093/ije/dyl170

Malmberg, M., Overbeek, G., Monshouwer, K., Lammers, J., Vollebergh, W. A., & Engels, R. C. (2010). Substance use risk profiles and associations with early substance use in adolescence. *Journal of Behavioral Medicine*, *33*(6), 474–485.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., McCarthy, M. I., Ramos, E. M., Cardon, L. R., Chakravarti, A., Cho, J. H., Guttmacher, A. E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C. N., Slatkin, M., Valle, D., Whittemore, A. S., … Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*(7265), 747–753. https://doi.org/10.1038/nature08494

Manzano, C., Hernández Castellano, M., Roman, L., Astals, M., Compta, B., & others. (2016). Maternal smoking during pregnancy and its impact on postnatal neurodevelopment. *Clinics Mother Child Health*, *13*(249), 2.

Marees, A. T., Kluiver, H. de, Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., & Derks, E. M. (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International Journal of Methods in Psychiatric Research*, *27*(2), e1608. https://doi.org/10.1002/mpr.1608

Markunas, C. A., Wilcox, A. J., Xu, Z., Joubert, B. R., Harlid, S., Panduri, V., Håberg, S. E., Nystad, W., London, S. J., Sandler, D. P., Lie, R. T., Wade, P. A., & Taylor, J. A. (2016). Maternal Age at Delivery Is Associated with an Epigenetic Signature in Both Newborns and Adults. *PLoS ONE*, *11*(7). https://doi.org/10.1371/journal.pone.0156361

Martin, C. L., Jima, D., Sharp, G. C., McCullough, L. E., Park, S. S., Gowdy, K. M., Skaar, D., Cowley, M., Maguire, R. L., Fuemmeler, B., Collier, D., Relton, C. L., Murphy, S. K., & Hoyo, C. (2019). Maternal pre-pregnancy obesity, offspring cord blood DNA methylation, and offspring cardiometabolic health in early childhood: An epigenome-wide association

study. *Epigenetics*, *14*(4), 325–340. https://doi.org/10.1080/15592294.2019.1581594

Marshall, O., Blaylock, R., Murphy, C., & Sanders, J. (2021). *Risk messages relating to fertility and pregnancy: A media content analysis* (6:114). Wellcome Open Research. https://doi.org/10.12688/wellcomeopenres.16744.1

Martin, J., Tilling, K., Hubbard, L., Stergiakouli, E., Thapar, A., Davey Smith, G., O'Donovan, M. C., & Zammit, S. (2016). Association of Genetic Risk for Schizophrenia With Nonparticipation Over Time in a Population-Based Cohort Study. *American Journal of Epidemiology*, *183*(12), 1149–1158. https://doi.org/10.1093/aje/kww009

Martin, A. R., Daly, M. J., Robinson, E. B., Hyman, S. E., & Neale, B. M. (2019). Predicting polygenic risk of psychiatric disorders. *Biological Psychiatry*, *86*(2), 97–109. https://doi.org/10.1016/j.biopsych.2018.12.015

Martino, D. J., Tulic, M. K., Gordon, L., Hodder, M., Richman, T. R., Metcalfe, J., Prescott, S. L., & Saffery, R. (2011). Evidence for age-related and individual-specific changes in DNA methylation profile of mononuclear cells during early immune development in humans. *Epigenetics*, *6*(9), 1085–1094. https://doi.org/10.4161/epi.6.9.16401

Marufu, T. C., Ahankari, A., Coleman, T., & Lewis, S. (2015). Maternal smoking and the risk of still birth: Systematic review and meta-analysis. *BMC Public Health*, *15*. https://doi.org/10.1186/s12889-015-1552-5

Massey, S. H., & Compton, M. T. (2013). Psychological Differences Between Smokers Who Spontaneously Quit During Pregnancy and Those Who Do Not: A Review of Observational Studies and Directions for Future Research. *Nicotine & Tobacco Research*, *15*(2), 307–319. https://doi.org/10.1093/ntr/nts142

Matijasevich, A., Murray, J., Cooper, P. J., Anselmi, L., Barros, A. J. D., Barros, F. C., & Santos, I. S. (2015). Trajectories of maternal depression and offspring psychopathology at 6 years: 2004 Pelotas cohort study. *Journal of Affective Disorders*, *174*, 424–431. https://doi.org/10.1016/j.jad.2014.12.012

Maughan, B., Taylor, A., Caspi, A., & Moffitt, T. E. (2004). Prenatal Smoking and Early Childhood Conduct Problems: Testing Genetic and

Environmental Explanations of the Association. *Archives of General Psychiatry*, *61*(8), 836–843. https://doi.org/10.1001/archpsyc.61.8.836

Maurice-Stam, H., Haverman, L., Splinter, A., van Oers, H. A., Schepers, S. A., & Grootenhuis, M. A. (2018). Dutch norms for the Strengths and Difficulties Questionnaire (SDQ) – parent form for children aged 2–18 years. *Health and Quality of Life Outcomes*, *16*(1). https://doi.org/10.1186/s12955-018-0948-1

McCartney, D. L., Hillary, R. F., Stevenson, A. J., Ritchie, S. J., Walker, R. M., Zhang, Q., Morris, S. W., Bermingham, M. L., Campbell, A., Murray, A. D., Whalley, H. C., Gale, C. R., Porteous, D. J., Haley, C. S., McRae, A. F., Wray, N. R., Visscher, P. M., McIntosh, A. M., Evans, K. L., … Marioni, R. E. (2018). Epigenetic prediction of complex traits and death. *Genome Biology*, *19*(1), 136. https://doi.org/10.1186/s13059-018-1514-1

McDonnell, B. P., & Regan, C. (2019). Smoking in pregnancy: Pathophysiology of harm and current evidence for monitoring and cessation. *The Obstetrician & Gynaecologist*, *21*(3), 169–175. https://doi.org/10.1111/tog.12585

McGowan, P. O., Sasaki, A., D'Alessio, A. C., Dymov, S., Labonté, B., Szyf, M., Turecki, G., & Meaney, M. J. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience*, *12*(3), 342–348. https://doi.org/10.1038/nn.2270

McGregor, K., Bernatsky, S., Colmegna, I., Hudson, M., Pastinen, T., Labbe, A., & Greenwood, C. M. T. (2016). An evaluation of methods correcting for cell-type heterogeneity in DNA methylation studies. *Genome Biology*, *17*(1), 84. https://doi.org/10.1186/s13059-016-0935-y

Meltzer, H., Gatward, R., Goodman, R., Ford, T. (2000). *The mental health of children and adolescents in Great Britain*. The Stationery Office. https://www.dawba.info/abstracts/B-CAMHS99_original_survey_report.pdf

Menezes, A. M. B., Murray, J., László, M., Wehrmeister, F. C., Hallal, P. C., Gonçalves, H., Assunção, M. C. F., Menezes, C. B., & Barros, F. C. (2013). Happiness and Depression in Adolescence after Maternal Smoking during Pregnancy: Birth Cohort Study. *PLOS ONE*, *8*(11), e80370. https://doi.org/10.1371/journal.pone.0080370

Merid, S. K., Novoloaca, A., Sharp, G. C., Küpers, L. K., Kho, A. T., Roy, R., Gao, L., Annesi-Maesano, I., Jain, P., Plusquin, M., Kogevinas, M., Allard, C., Vehmeijer, F. O., Kazmi, N., Salas, L. A., Rezwan, F. I., Zhang, H., Sebert, S., Czamara, D., … Melén, E. (2020). Epigenome-wide meta-analysis of blood DNA methylation in newborns and children identifies numerous loci related to gestational age. *Genome Medicine*, *12*(1), 25. https://doi.org/10.1186/s13073-020-0716-9

Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime Prevalence of Mental Disorders in U.S. Adolescents: Results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(10), 980–989. https://doi.org/10.1016/j.jaac.2010.05.017

Middeldorp, C. M., Felix, J. F., Mahajan, A., Ahluwalia, T. S., Auvinen, J., Bartels, M., Bilbao, J. R., Bisgaard, H., Bønnelykke, K., Boomsma, D. I., Bradfield, J. P., Bustamante, M., Chen, Z., Curtin, J. A., Custovic, A., Smith, G. D., Davies, G. E., Duijts, L., Eastwood, P. R., … Early Growth Genetics (EGG) consortium. (2019). The Early Growth Genetics (EGG) and EArly Genetics and Lifecourse Epidemiology (EAGLE) consortia: Design, results and future prospects. *European Journal of Epidemiology*, *34*(3), 279–300. https://doi.org/10.1007/s10654-019-00502-9

Miguel, P. M., Pereira, L. O., Silveira, P. P., & Meaney, M. J. (2019). Early environmental influences on the development of children’s brain structure and function. *Developmental Medicine & Child Neurology*, *61*(10), 1127–1133. https://doi.org/10.1111/dmcn.14182

Mikkelsen, S., Olsen, J., Niclasen, J., & Bech, B. (2017). Maternal Caffeine Consumption during Pregnancy and Behavioral Disorders in 11-Year-Old Offspring: A Danish National Birth Cohort Study. *The Journal of Pediatrics*, *189*. https://doi.org/10.1016/j.jpeds.2017.06.051

Millard, L. A. C., Davies, N. M., Timpson, N. J., Tilling, K., Flach, P. A., & Smith, G. D. (2015). MR-PheWAS: Hypothesis prioritization among potential causal effects of body mass index on many outcomes, using Mendelian randomization. *Scientific Reports*, *5*(1), 16645. https://doi.org/10.1038/srep16645

Miller, D. B., & O'Callaghan, J. P. (2002). Neuroendocrine aspects of the response to stress. *Metabolism*, *51*(6), 5–10. https://doi.org/10.1053/meta.2002.33184

Mills, M. C., & Rahal, C. (2019). A scientometric review of genome-wide association studies. *Communications Biology*, *2*(1), 1–11. https://doi.org/10.1038/s42003-018-0261-x

Mills, M. C., & Tropf, F. C. (2020). Sociology, Genetics, and the Coming of Age of Sociogenomics. *Annual Review of Sociology*, *46*(1), 553–581. https://doi.org/10.1146/annurev-soc-121919-054756

Min, J. L., Hemani, G., Davey Smith, G., Relton, C., & Suderman, M. (2018). Meffil: Efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics (Oxford, England)*, *34*(23), 3983–3989. https://doi.org/10.1093/bioinformatics/bty476

Min, J. L., Hemani, G., Hannon, E., Dekkers, K. F., Castillo-Fernandez, J., Luijk, R., Carnero-Montoro, E., Lawson, D. J., Burrows, K., Suderman, M., Breterick, A. D., Richardson, T. G., Klughammer, J., Iotchkova, V., Sharp, G. C., Khleifat, A. A., Shatunov, A., Iacoangeli, A., McArdle, W. L., … Relton, C. L. (2020). Genomic and phenomic insights from an atlas of genetic effects on DNA methylation. *MedRxiv*, 2020.09.01.20180406. https://doi.org/10.1101/2020.09.01.20180406

Miyake, Y., Tanaka, K., Okubo, H., Sasaki, S., & Arakawa, M. (2019). Maternal caffeine intake in pregnancy is inversely related to childhood peer problems in Japan: The Kyushu Okinawa Maternal and Child Health Study. *Nutritional Neuroscience*, *22*(11), 817–824. https://doi.org/10.1080/1028415X.2018.1450089

Mooney, M. A., Ryabinin, P., Wilmot, B., Bhatt, P., Mill, J., & Nigg, J. T. (2020). Large epigenome-wide association study of childhood ADHD identifies peripheral DNA methylation associated with disease and polygenic risk burden. *Translational Psychiatry*, *10*(1), 1–12. https://doi.org/10.1038/s41398-020-0710-4

Moran, S., Arribas, C., & Esteller, M. (2016). Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics*, *8*(3), 389–399. https://doi.org/10.2217/epi.15.114

Morgan, A. J., & Jorm, A. F. (2008). Self-help interventions for depressive disorders and depressive symptoms: A systematic review. *Annals of General Psychiatry*, *7*(1), 13. https://doi.org/10.1186/1744-859X-7-13

Morin, A. J. S., Arens, A. K., Maïano, C., Ciarrochi, J., Tracey, D., Parker, P. D., & Craven, R. G. (2017). Reciprocal Relationships between Teacher Ratings of Internalizing and Externalizing Behaviors in Adolescents with Different Levels of Cognitive Abilities. *Journal of Youth and Adolescence*, *46*(4), 801–825. https://doi.org/10.1007/s10964-016-0574-3

Mourady, D., Richa, S., Karam, R., Papazian, T., Moussa, F. H., Osta, N. E., Kesrouani, A., Azouri, J., Jabbour, H., Hajj, A., & Khabbaz, L. R. (2017). Associations between quality of life, physical activity, worry, depression and insomnia: A cross-sectional designed study in healthy pregnant women. *PLOS ONE*, *12*(5), e0178181. https://doi.org/10.1371/journal.pone.0178181

Moylan, S., Gustavson, K., Øverland, S., Karevold, E. B., Jacka, F. N., Pasco, J. A., & Berk, M. (2015). The impact of maternal smoking during pregnancy on depressive and anxiety behaviors in children: The Norwegian Mother and Child Cohort Study. *BMC Medicine*, *13*(1), 24. https://doi.org/10.1186/s12916-014-0257-4

Mulder, R. H., Neumann, A., Cecil, C. A. M., Walton, E., Houtepen, L. C., Simpkin, A. J., Rijlaarsdam, J., Heijmans, B. T., Gaunt, T. R., Felix, J. F., Jaddoe, V. W. V., Bakermans-Kranenburg, M. J., Tiemeier, H., Relton, C. L., IJzendoorn, M. H. van, & Suderman, M. (2020). Epigenome-wide change and variation in DNA methylation from birth to late adolescence. *BioRxiv*, 2020.06.09.142620. https://doi.org/10.1101/2020.06.09.142620

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: When selection bias can substantially influence observed associations. *International Journal of Epidemiology*, *47*(1), 226–235. https://doi.org/10.1093/ije/dyx206

Murgatroyd, C., & Spengler, D. (2011). Epigenetic programming of the HPA axis: Early life decides. *Stress*, *14*(6), 581–589. https://doi.org/10.3109/10253890.2011.602146

Murphy, C., Brown, T., Trickey, H., Sanders, J., Blaylock, R., Dean, C., Hennsessy, M., Schellhas, L., Sharp, G. C., Zuccolo, L., Munafò, M. R.,

Cairns, K., Booker, M., Tennant, P., Fisher, J., Lee, E., Williams, E., Duff, E., Petersen, I., & Marshall, E. (2020, September 3). *It remains unclear whether caffeine causes adverse pregnancy outcomes; but naive policy recommendations could cause harm [Letter to the editor].* 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

Murray, G. K., Lin, T., Austin, J., McGrath, J. J., Hickie, I. B., & Wray, N. R. (2021). Could Polygenic Risk Scores Be Useful in Psychiatry?: A Review. *JAMA Psychiatry*, *78*(2), 210–219. https://doi.org/10.1001/jamapsychiatry.2020.3042

Najman, J. M., Williams, G. M., Nikles, J., Spence, S., Bor, W., O'Callaghan, M., Le Brocque, R., Andersen, M. J., & Shuttlewood, G. J. (2001). Bias influencing maternal reports of child behaviour and emotional state. *Social Psychiatry and Psychiatric Epidemiology*, *36*(4), 186–194. https://doi.org/10.1007/s001270170062

Neumann, A., Walton, E., Alemany, S., Cecil, C., González, J. R., Jima, D. D., Lahti, J., Tuominen, S. T., Barker, E. D., Binder, E., Caramaschi, D., Carracedo, Á., Czamara, D., Evandt, J., Felix, J. F., Fuemmeler, B. F., Gutzkow, K. B., Hoyo, C., Julvez, J., … Tiemeier, H. (2020). Association between DNA methylation and ADHD symptoms from birth to school age: A prospective meta-analysis. *Translational Psychiatry*, *10*(1), 1–11. https://doi.org/10.1038/s41398-020-01058-z

Nicolopoulos, K., Mulugeta, A., Zhou, A., & Hyppönen, E. (2020). Association between habitual coffee consumption and multiple disease outcomes: A Mendelian randomisation phenome-wide association study in the UK Biobank. *Clinical Nutrition*, *39*(11), 3467–3476. https://doi.org/10.1016/j.clnu.2020.03.009

Nivard, M. G., Gage, S. H., Hottenga, J. J., van Beijsterveldt, C. E. M., Abdellaoui, A., Bartels, M., Baselmans, B. M. L., Ligthart, L., Pourcain, B. S., Boomsma, D. I., Munafò, M. R., & Middeldorp, C. M. (2017). Genetic Overlap Between Schizophrenia and Developmental

Psychopathology: Longitudinal and Multivariate Polygenic Risk Prediction of Common Psychiatric Traits During Development. *Schizophrenia Bulletin*, *43*(6), 1197–1207. https://doi.org/10.1093/schbul/sbx031

No safe level of coffee drinking for pregnant women, study says. (2020, August 24). The Guardian. http://www.theguardian.com/lifeandstyle/2020/aug/24/no-safe-level-coffee-drinking-pregnant-women-study-says

Nomura, Y., Marks, D. J., & Halperin, J. M. (2010). Prenatal Exposure to Maternal and Paternal Smoking on Attention Deficit Hyperactivity Disorders Symptoms and Diagnosis in Offspring. *The Journal of Nervous and Mental Disease*, *198*(9), 672–678. https://doi.org/10.1097/NMD.0b013e3181ef3489

Oberlander, T. F., Papsdorf, M., Brain, U. M., Misri, S., Ross, C., & Grunau, R. E. (2010). Prenatal effects of selective serotonin reuptake inhibitor antidepressants, serotonin transporter promoter genotype (SLC6A4), and maternal mood on child behavior at 3 years of age. *Archives of Pediatrics & Adolescent Medicine*, *164*(5), 444–451. https://doi.org/10.1001/archpediatrics.2010.51

Oberlander, T. F., Weinberg, J., Papsdorf, M., Grunau, R., Misri, S., & Devlin, A. M. (2008). Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics*, *3*(2), 97–106. https://doi.org/10.4161/epi.3.2.6034

O’Donnell, K. J., & Meaney, M. J. (2017). Fetal Origins of Mental Health: The Developmental Origins of Health and Disease Hypothesis. *American Journal of Psychiatry*, *174*(4), 319–328. https://doi.org/10.1176/appi.ajp.2016.16020138

O’hara, M. W., & Swain, A. M. (1996). Rates and risk of postpartum depression—A meta-analysis. *International Review of Psychiatry*, *8*(1), 37–54. https://doi.org/10.3109/09540269609037816

Olsson, C. A., Foley, D. L., Parkinson-Bates, M., Byrnes, G., McKenzie, M., Patton, G. C., Morley, R., Anney, R. J. L., Craig, J. M., & Saffery, R. (2010). Prospects for epigenetic research within cohort studies of psychological disorder: A pilot investigation of a peripheral cell marker of

epigenetic risk for depression. *Biological Psychology*, *83*(2), 159–165. https://doi.org/10.1016/j.biopsycho.2009.12.003

Pariante, C. M., & Lightman, S. L. (2008). The HPA axis in major depression: Classical theories and new developments. *Trends in Neurosciences*, *31*(9), 464–468. https://doi.org/10.1016/j.tins.2008.06.006

Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). *Mental health of young people: A global public-health challenge*. *369*, 12.

Patti, M. A., Li, N., Eliot, M., Newschaffer, C., Yolton, K., Khoury, J., Chen, A., Lanphear, B. P., Lyall, K., Hertz-Picciotto, I., Fallin, M. D., Croen, L. A., & Braun, J. M. (2021). Association between self-reported caffeine intake during pregnancy and social responsiveness scores in childhood: The EARLI and HOME studies. *PLOS ONE*, *16*(1), e0245079. https://doi.org/10.1371/journal.pone.0245079

Peck, L., Borle, K., Folkersen, L., & Austin, J. (2021). Why do people seek out polygenic risk scores for complex disorders, and how do they understand and react to results?. *European journal of human genetics : EJHG*, 10.1038/s41431-021-00929-3. Advance online publication. https://doi.org/10.1038/s41431-021-00929-3

Pedersen, M. U., Thomsen, K. R., Heradstveit, O., Skogen, J. C., Hesse, M., & Jones, S. (2018). Externalizing behavior problems are related to substance use in adolescents across six samples from Nordic countries. *European Child & Adolescent Psychiatry*, *27*(12), 1551–1561. https://doi.org/10.1007/s00787-018-1148-6

Pembrey, M. (2004). The Avon Longitudinal Study of Parents and Children (ALSPAC): A resource for genetic epidemiology. *European Journal of Endocrinology*, *151*(Suppl_3), U125–U129. https://doi.org/10.1530/eje.0.151u125

Perera, F., & Herbstman, J. (2011). Prenatal environmental exposures, epigenetics, and disease. *Reproductive Toxicology*, *31*(3), 363–373. https://doi.org/10.1016/j.reprotox.2010.12.055

Phipson, B., Maksimovic, J., & Oshlack, A. (2015). missMethyl: An R package for analysing methylation data from Illuminas HumanMethylation450 platform. *Bioinformatics*, btv560.

Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy,

P., Van Djik, S., Muhlhausler, B., Stirzaker, C., & Clark, S. J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology*, *17*(1), 208. https://doi.org/10.1186/s13059-016-1066-1

Pingault, J.-B., Cecil, C. A. M., Murray, J., Munafò, M. R., & Viding, E. (2017). Causal Inference in Psychopathology: A Systematic Review of Mendelian Randomisation Studies Aiming to Identify Environmental Risk Factors for Psychopathology. *Psychopathology Review*, *a4*(1), 4–25. https://doi.org/10.5127/pr.038115

Pingault, J.-B., O'Reilly, P. F., Schoeler, T., Ploubidis, G. B., Rijsdijk, F., & Dudbridge, F. (2018). Using genetic data to strengthen causal inference in observational research. *Nature Reviews Genetics*, *19*(9), 566–580. https://doi.org/10.1038/s41576-018-0020-3

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709. https://doi.org/10.1038/ng.3285

Polimanti, R., Kranzler, H. R., & Gelernter, J. (2016). Phenome-Wide Association Study for Alcohol and Nicotine Risk Alleles in 26394 Women. *Neuropsychopharmacology*, *41*(11), 2688–2696. https://doi.org/10.1038/npp.2016.72

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Pritchard, C. W. (1994). Depression and smoking in pregnancy in Scotland. *Journal of Epidemiology and Community Health*, *48*(4), 377–382.

Purves, K. L., Coleman, J. R. I., Meier, S. M., Rayner, C., Davis, K. A. S., Cheesman, R., Bækvad-Hansen, M., Børglum, A. D., Wan Cho, S., Jürgen Deckert, J., Gaspar, H. A., Bybjerg-Grauholm, J., Hettema, J. M., Hotopf, M., Hougaard, D., Hübel, C., Kan, C., McIntosh, A. M., Mors, O., … Eley, T. C. (2020). A major role for common genetic variation in anxiety disorders. *Molecular Psychiatry*, *25*(12), 3292–3303. https://doi.org/10.1038/s41380-019-0559-1

Qian, J., Chen, Q., Ward, S. M., Duan, E., & Zhang, Y. (2020). Impacts of Caffeine during Pregnancy. *Trends in Endocrinology & Metabolism*, *31*(3), 218–227. https://doi.org/10.1016/j.tem.2019.11.004

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Ramsawh, H. J., Weisberg, R. B., Dyck, I., Stout, R., & Keller, M. B. (2011). Age of onset, clinical characteristics, and 15-year course of anxiety disorders in a prospective, longitudinal, observational study. *Journal of Affective Disorders*, *132*(1–2), 260–264. https://doi.org/10.1016/j.jad.2011.01.006

Rauschert, S., Melton, P. E., Heiskala, A., Karhunen, V., Burdge, G., Craig, J. M., Godfrey, K. M., Lillycrop, K., Mori, T. A., Beilin, L. J., Oddy, W. H., Pennell, C., Järvelin, M.-R., Sebert, S., & Huang, R.-C. (2020). Machine Learning-Based DNA Methylation Score for Fetal Exposure to Maternal Smoking: Development and Validation in Samples Collected from Adolescents and Adults. *Environmental Health Perspectives*, *128*(9), 97003. https://doi.org/10.1289/EHP6076

Reese, S. E., Zhao, S., Wu, M. C., Joubert, B. R., Parr, C. L., Håberg, S. E., Ueland, P. M., Nilsen, R. M., Midttun, Ø., Vollset, S. E., Peddada, S. D., Nystad, W., & London, S. J. (2017). DNA Methylation Score as a Biomarker in Newborns for Sustained Maternal Smoking during Pregnancy. *Environmental Health Perspectives*, *125*(4), 760–766. https://doi.org/10.1289/EHP333

Reinius, L. E., Acevedo, N., Joerink, M., Pershagen, G., Dahlén, S.-E., Greco, D., Söderhäll, C., Scheynius, A., & Kere, J. (2012). Differential DNA Methylation in Purified Human Blood Cells: Implications for Cell Lineage and Studies on Disease Susceptibility. *PLoS ONE*, *7*(7), e41361. https://doi.org/10.1371/journal.pone.0041361

Relton, C. L., & Davey Smith, G. (2012). Two-step epigenetic Mendelian randomization: A strategy for establishing the causal role of epigenetic processes in pathways to disease. *International Journal of Epidemiology*, *41*(1), 161–176. https://doi.org/10.1093/ije/dyr233

Relton, C. L., Gaunt, T., McArdle, W., Ho, K., Duggirala, A., Shihab, H., Woodward, G., Lyttleton, O., Evans, D. M., Reik, W., Paul, Y.-L., Ficz, G., Ozanne, S. E., Wipat, A., Flanagan, K., Lister, A., Heijmans, B. T.,

Ring, S. M., & Davey Smith, G. (2015). Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *International Journal of Epidemiology*, *44*(4), 1181–1190. https://doi.org/10.1093/ije/dyv072

Relton, C. L., Hartwig, F. P., & Davey Smith, G. (2015). From stem cells to the law courts: DNA methylation, the forensic epigenome and the possibility of a biosocial archive. *International Journal of Epidemiology*, *44*(4), 1083–1093. https://doi.org/10.1093/ije/dyv198

Reyes, C. M., & Cornelis, M. (2018). Caffeine in the diet: Country-level consumption and guidelines. *Nutrients*, *10*(11), 1772.

Reynolds, C. A., Barlow, T., & Pedersen, N. L. (2006). Alcohol, tobacco and caffeine use: Spouse similarity processes. *Behavior Genetics*, *36*(2), 201–215. https://doi.org/10.1007/s10519-005-9026-7

Reynolds, C. A., Tan, Q., Munoz, E., Jylhävä, J., Hjelmborg, J., Christiansen, L., Hägg, S., & Pedersen, N. L. (2020). A decade of epigenetic change in aging twins: Genetic and environmental contributions to longitudinal DNA methylation. *Aging Cell*, *19*(8), e13197. https://doi.org/10.1111/acel.13197

Riaz, M., Lewis, S., Naughton, F., & Ussher, M. (2018). Predictors of smoking cessation during pregnancy: A systematic review and meta-analysis. *Addiction*, *113*(4), 610–622. https://doi.org/10.1111/add.14135

Rice, F., Langley, K., Woodford, C., Smith, G. D., & Thapar, A. (2018). Identifying the contribution of prenatal risk factors to offspring development and psychopathology: What designs to use and a critique of literature on maternal smoking and stress in pregnancy. *Development and Psychopathology*, *30*(3), 1107–1128.

Richardson, T. G., Richmond, R. C., North, T.-L., Hemani, G., Davey Smith, G., Sharp, G. C., & Relton, C. L. (2019). An integrative approach to detect epigenetic mechanisms that putatively mediate the influence of lifestyle exposures on disease susceptibility. *International Journal of Epidemiology*, *48*(3), 887–898. https://doi.org/10.1093/ije/dyz119

Richmond, R. C., Al-Amin, A., Davey Smith, G., & Relton, C. L. (2014). Approaches for drawing causal inferences from epidemiological birth cohorts: A Review. *Early Human Development*, *90*(11), 769–780.

https://doi.org/10.1016/j.earlhumdev.2014.08.023

Richmond, R. C., Suderman, M., Langdon, R., Relton, C. L., & Davey Smith, G. (2018). DNA methylation as a marker for prenatal smoke exposure in adults. *International Journal of Epidemiology*, *47*(4), 1120–1130. https://doi.org/10.1093/ije/dyy091

Richmond, R. C., Timpson, N. J., Felix, J. F., Palmer, T., Gaillard, R., McMahon, G., Davey Smith, G., Jaddoe, V. W., & Lawlor, D. A. (2017). Using Genetic Variation to Explore the Causal Effect of Maternal Pregnancy Adiposity on Future Offspring Adiposity: A Mendelian Randomisation Study. *PLoS Medicine*, *14*(1). https://doi.org/10.1371/journal.pmed.1002221

Riglin, L., Collishaw, S., Richards, A., Thapar, A. K., Rice, F., Maughan, B., O'Donovan, M. C., & Thapar, A. (2018). The impact of schizophrenia and mood disorder risk alleles on emotional problems: Investigating change from childhood to middle age. *Psychological Medicine*, *48*(13), 2153–2158. https://doi.org/10.1017/S0033291717003634

Ringoot, A. P., Tiemeier, H., Jaddoe, V. W. V., So, P., Hofman, A., Verhulst, F. C., & Jansen, P. W. (2015). Parental depression and child well-being: Young children's self-reports helped addressing biases in parent reports. *Journal of Clinical Epidemiology*, *68*(8), 928–938. https://doi.org/10.1016/j.jclinepi.2015.03.009

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47.

Roberts, S., Lester, K. J., Hudson, J. L., Rapee, R. M., Creswell, C., Cooper, P. J., Thirlwall, K. J., Coleman, J. R. I., Breen, G., Wong, C. C. Y., & Eley, T. C. (2014). Serotonin tranporter methylation and response to cognitive behaviour therapy in children with anxiety disorders. *Translational Psychiatry*, *4*(9), e444–e444. https://doi.org/10.1038/tp.2014.83

Robinson, M., McLean, N. J., Oddy, W. H., Mattes, E., Bulsara, M., Li, J., Zubrick, S. R., Stanley, F. J., & Newnham, J. P. (2010). Smoking cessation in pregnancy and the risk of child behavioural problems: A longitudinal prospective cohort study. *Journal of Epidemiology & Community Health*, *64*(7), 622–629.

https://doi.org/10.1136/jech.2009.088658

Román-Gálvez, R. M., Amezcua-Prieto, C., Olmedo-Requena, R., Saad, A. L.-M., Martínez-Galiano, J. M., & Bueno-Cavanillas, A. (2018). Partner smoking influences whether mothers quit smoking during pregnancy: A prospective cohort study. *BJOG: An International Journal of Obstetrics & Gynaecology*, *125*(7), 820–827. https://doi.org/10.1111/1471-0528.14986

Ross, E. J., Graham, D. L., Money, K. M., & Stanwood, G. D. (2015). Developmental Consequences of Fetal Exposure to Drugs: What We Know and What We Still Must Learn. *Neuropsychopharmacology*, *40*(1), 61–87. https://doi.org/10.1038/npp.2014.147

Royston, P., Altman, D. G., & Sauerbrei, W. (2006). Dichotomizing continuous predictors in multiple regression: A bad idea. *Statistics in Medicine*, *25*(1), 127–141. https://doi.org/10.1002/sim.2331

Roza, S. J., Verhulst, F. C., Jaddoe, V. W., Steegers, E. A., Mackenbach, J. P., Hofman, A., & Tiemeier, H. (2009). Maternal smoking during pregnancy and child behaviour problems: The Generation R Study. *International Journal of Epidemiology*, *38*(3), 680–689. https://doi.org/10.1093/ije/dyn163

Sadler, K., Vizard, T., Ford, T., Marchesell, F., Pearce, N., Mandalia, D., Davis, J., Brodie, E., Forbes, N., Goodman, A., & others. (2018). *Mental health of children and young people in England, 2017*.

Saffari, A., Silver, M. J., Zavattari, P., Moi, L., Columbano, A., Meaburn, E. L., & Dudbridge, F. (2018). Estimation of a significance threshold for epigenome-wide association studies. *Genetic Epidemiology*, *42*(1), 20–33. https://doi.org/10.1002/gepi.22086

Sala, M., Perez, J., Soloff, P., Ucelli di Nemi, S., Caverzasi, E., Soares, J. C., & Brambilla, P. (2004). Stress and hippocampal abnormalities in psychiatric disorders. *European Neuropsychopharmacology*, *14*(5), 393–405. https://doi.org/10.1016/j.euroneuro.2003.12.005

Salihu, H. M., & Wilson, R. E. (2007). Epidemiology of prenatal smoking and perinatal outcomes. *Early Human Development*, *83*(11), 713–720. https://doi.org/10.1016/j.earlhumdev.2007.08.002

Sallis, H., Szekely, E., Neumann, A., Jolicoeur-Martineau, A., van IJzendoorn,

M., Hillegers, M., Greenwood, C. M. T., Meaney, M. J., Steiner, M., Tiemeier, H., Wazana, A., Pearson, R. M., & Evans, J. (2019). General psychopathology, internalising and externalising in children and functional outcomes in late adolescence. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *60*(11), 1183–1190. https://doi.org/10.1111/jcpp.13067

Schellhas, L., Haan, E., Easey, K. E., Wootton, R. E., Sallis, H. M., Sharp, G. C., Munafò, M. R., & Zuccolo, L. (2021). Maternal and child genetic liability for smoking and caffeine consumption and child mental health: An intergenerational genetic risk score analysis in the ALSPAC cohort. *Addiction*. https://doi.org/10.1111/add.15521

Schneider, S., Huy, C., Schütz, J., & Diehl, K. (2010). Smoking cessation during pregnancy: A systematic literature review. *Drug and Alcohol Review*, *29*(1), 81–90. https://doi.org/10.1111/j.1465-3362.2009.00098.x

Schneider, S., & Schütz, J. (2008). Who smokes during pregnancy? A systematic literature review of population-based surveys conducted in developed countries between 1997 and 2006. *The European Journal of Contraception & Reproductive Health Care*, *13*(2), 138–147. https://doi.org/10.1080/13625180802027993

Schreiber, G. B., Maffeo, C. E., Robins, M., Masters, M. N., & Bond, A. P. (1988). Measurement of coffee and caffeine intake: Implications for epidemiologic research. *Preventive Medicine*, *17*(3), 280–294. https://doi.org/10.1016/0091-7435(88)90004-7

Sengpiel, V., Elind, E., Bacelis, J., Nilsson, S., Grove, J., Myhre, R., Haugen, M., Meltzer, H. M., Alexander, J., Jacobsson, B., & Brantsæter, A.-L. (2013). 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*, *11*(1), 42. https://doi.org/10.1186/1741-7015-11-42

Seo, D., Patrick, C. J., & Kennealy, P. J. (2008). Role of Serotonin and Dopamine System Interactions in the Neurobiology of Impulsive Aggression and its Comorbidity with other Clinical Disorders. *Aggression and Violent Behavior*, *13*(5), 383–395. https://doi.org/10.1016/j.avb.2008.06.003

Shaffer, J. P. (1995). Multiple Hypothesis Testing. *Annual Review of Psychology*,

46(1), 561–584. https://doi.org/10.1146/annurev.ps.46.020195.003021

Shabalin, A. A., & Aberg, K. A. (2015). Candidate gene methylation studies are at high risk of erroneous conclusions. *Epigenomics*, *7*(1), 13–15. https://doi.org/10.2217/epi.14.70

Shah, P. S. (2010). Parity and low birth weight and preterm birth: A systematic review and meta-analyses. *Acta Obstetricia et Gynecologica Scandinavica*, *89*(7), 862–875. https://doi.org/10.3109/00016349.2010.486827

Sharp, G. C., Alfano, R., Ghantous, A., Urquiza, J., Rifas-Shiman, S. L., Page, C. M., Jin, J., Fernández-Barrés, S., Santorelli, G., Tindula, G., & 36 other members of the Pregnancy and Childhood Epigenetics (PACE) consortium. (2021). Paternal body mass index and offspring DNA methylation: Findings from the PACE consortium. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyaa267

Sharp, G. C., Arathimos, R., Reese, S. E., Page, C. M., Felix, J., Küpers, L. K., Rifas-Shiman, S. L., Liu, C., The Cohorts for Heart and Aging Research in Genomic Epidemiology plus (CHARGE +) methylation alcohol working group, Burrows, K., Zhao, S., Magnus, M. C., Duijts, L., Corpeleijn, E., DeMeo, D. L., Litonjua, A., Baccarelli, A., Hivert, M.-F., Oken, E., … Zuccolo, L. (2018). Maternal alcohol consumption and offspring DNA methylation: Findings from six general population-based birth cohorts. *Epigenomics*, *10*(1), 27–42. https://doi.org/10.2217/epi-2017-0095

Sharp, G. C., Lawlor, D. A., & Richardson, S. S. (2018). 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*, *213*, 20–27. https://doi.org/10.1016/j.socscimed.2018.07.035

Sharp, G. C., & Relton, C. L. (2017). Epigenetics and noncommunicable diseases. *Epigenomics*, *9*(6), 789–791. https://doi.org/10.2217/epi-2017-0045

Sharp, G. C., Salas, L. A., Monnereau, C., Allard, C., Yousefi, P., Everson, T. M., Bohlin, J., Xu, Z., Huang, R.-C., Reese, S. E., Xu, C.-J., Baïz, N., Hoyo, C., Agha, G., Roy, R., Holloway, J. W., Ghantous, A., Merid, S. K., Bakulski, K. M., … Relton, C. L. (2017). Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation: Findings

from the pregnancy and childhood epigenetics (PACE) consortium. *Human Molecular Genetics*, *26*(20), 4067–4085. https://doi.org/10.1093/hmg/ddx290

Sharp, G. C., Schellhas, L., Richardson, S. S., & Lawlor, D. A. (2019). Time to cut the cord: Recognizing and addressing the imbalance of DOHaD research towards the study of maternal pregnancy exposures. *Journal of Developmental Origins of Health and Disease*, *10*(05), 509–512. https://doi.org/10.1017/S2040174419000072

Shipton, D., Tappin, D. M., Vadiveloo, T., Crossley, J. A., Aitken, D. A., & Chalmers, J. (2009). Reliability of self reported smoking status by pregnant women for estimating smoking prevalence: A retrospective, cross sectional study. *The BMJ*, *339*. https://doi.org/10.1136/bmj.b4347

Silva, C. G., Metin, C., Fazeli, W., Machado, N. J., Darmopil, S., Launay, P.-S., Ghestem, A., Nesa, M.-P., Bassot, E., Szabo, E., Baqi, Y., Muller, C. E., Tome, A. R., Ivanov, A., Isbrandt, D., Zilberter, Y., Cunha, R. A., Esclapez, M., & Bernard, C. (2013). Adenosine Receptor Antagonists Including Caffeine Alter Fetal Brain Development in Mice. *Science Translational Medicine*, *5*(197), 197ra104-197ra104. https://doi.org/10.1126/scitranslmed.3006258

Smedberg, J., Lupattelli, A., Mårdby, A.-C., & Nordeng, H. (2014). Characteristics of women who continue smoking during pregnancy: A cross-sectional study of pregnant women and new mothers in 15 European countries. *BMC Pregnancy and Childbirth*, *14*(1), 213. https://doi.org/10.1186/1471-2393-14-213

Smith, A. M., Dwoskin, L. P., & Pauly, J. R. (2010). Early exposure to nicotine during critical periods of brain development: Mechanisms and consequences. *Journal of Pediatric Biochemistry*, *1*(2), 125–141. https://doi.org/10.3233/JPB-2010-0012

Staley, J. R., Blackshaw, J., Kamat, M. A., Ellis, S., Surendran, P., Sun, B. B., Paul, D. S., Freitag, D., Burgess, S., Danesh, J., Young, R., & Butterworth, A. S. (2016). PhenoScanner: A database of human genotype–phenotype associations. *Bioinformatics*, *32*(20), 3207–3209. https://doi.org/10.1093/bioinformatics/btw373

Starnawska, A., Tan, Q., Soerensen, M., McGue, M., Mors, O., Børglum, A. D.,

Christensen, K., Nyegaard, M., & Christiansen, L. (2019). Epigenome-wide association study of depression symptomatology in elderly monozygotic twins. *Translational Psychiatry*, *9*(1), 1–14. https://doi.org/10.1038/s41398-019-0548-9

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T. I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., & Lancet, D. (2016). The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current Protocols in Bioinformatics*, *54*(1). https://doi.org/10.1002/cpbi.5

Stephens, S. H., Hoft, N. R., Schlaepfer, I. R., Young, S. E., Corley, R. C., McQueen, M. B., Hopfer, C., Crowley, T., Stallings, M., Hewitt, J., & Ehringer, M. A. (2012). Externalizing Behaviors are associated with SNPs in the CHRNA5/CHRNA3/CHRNB4 gene cluster. *Behavior Genetics*, *42*(3), 402–414. https://doi.org/10.1007/s10519-011-9514-x

Stratil, J. M., Paudel, D., Setty, K. E., Menezes de Rezende, C. E., Monroe, A. A., Osuret, J., Scheel, I. B., Wildner, M., & Rehfuess, E. A. (2020). Advancing the WHO-INTEGRATE Framework as a Tool for Evidence-Informed, Deliberative Decision-Making Processes: Exploring the Views of Developers and Users of WHO Guidelines. *International Journal of Health Policy and Management*, 1. https://doi.org/10.34172/ijhpm.2020.193

Suderman, M., Simpkin, A., Sharp, G. C., Gaunt, T., Lyttleton, O., McArdle, W., Ring, S., Smith, G. D., & Relton, C. L. (2017). Sex-associated autosomal DNA methylation differences are wide-spread and stable throughout childhood. *BioRxiv*, 118265. https://doi.org/10.1101/118265

Suderman, M., Staley, J. R., French, R., Arathimos, R., Simpkin, A., & Tilling, K. (2018). dmrff: Identifying differentially methylated regions efficiently with power and control. *BioRxiv*. https://doi.org/10.1101/508556

Suderman, M., & Yousefi, P. (2020). *journalclub: Journal club papers*. https://github.com/perishky/journalclub

Swanson, J. A., Lee, J. W., & Hopp, J. W. (1994). Caffeine and nicotine: A review of their joint use and possible interactive effects in tobacco withdrawal. *Addictive Behaviors*, *19*(3), 229–256.

https://doi.org/10.1016/0306-4603(94)90027-2

Talens, R. P., Christensen, K., Putter, H., Willemsen, G., Christiansen, L., Kremer, D., Suchiman, H. E. D., Slagboom, P. E., Boomsma, D. I., & Heijmans, B. T. (2012). Epigenetic variation during the adult lifespan: Cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell*, *11*(4), 694–703. https://doi.org/10.1111/j.1474-9726.2012.00835.x

Tandon, M., Cardeli, E., & Luby, J. (2009). Internalizing Disorders in Early Childhood: A Review of Depressive and Anxiety Disorders. *Child and Adolescent Psychiatric Clinics of North America*, *18*(3), 593–610. https://doi.org/10.1016/j.chc.2009.03.004

Taylor, A. E., Carslake, D., de Mola, C. L., Rydell, M., Nilsen, T. I. L., Bjørngaard, J. H., Horta, B. L., Pearson, R., Rai, D., Galanti, M. R., Barros, F. C., Romundstad, P. R., Davey Smith, G., & Munafò, M. R. (2017). Maternal Smoking in Pregnancy and Offspring Depression: A cross cohort and negative control study. *Scientific Reports*, *7*(1), 12579. https://doi.org/10.1038/s41598-017-11836-3

Taylor, A. E., Davey Smith, G., Bares, C. B., Edwards, A. C., & Munafò, M. R. (2014). Partner smoking and maternal cotinine during pregnancy: Implications for negative control methods. *Drug and Alcohol Dependence*, *139*, 159–163. https://doi.org/10.1016/j.drugalcdep.2014.03.012

Taylor, A. E., Davey Smith, G., & Munafò, M. R. (2018). Associations of coffee genetic risk scores with consumption of coffee, tea and other beverages in the UK Biobank: Coffee genetic risk scores. *Addiction*, *113*(1), 148–157. https://doi.org/10.1111/add.13975

Taylor, A. E., Davies, N. M., Ware, J. J., VanderWeele, T., Smith, G. D., & Munafò, M. R. (2014). Mendelian randomization in health research: Using appropriate genetic variants and avoiding biased estimates. *Economics & Human Biology*, *13*, 99–106.

Taylor, A. E., Jones, H. J., Sallis, H., Euesden, J., Stergiakouli, E., Davies, N. M., Zammit, S., Lawlor, D. A., Munafò, M. R., Davey Smith, G., & Tilling, K. (2018). Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, *47*(4), 1207–1216. https://doi.org/10.1093/ije/dyy060

Teh, A. L., Pan, H., Chen, L., Ong, M.-L., Dogra, S., Wong, J., MacIsaac, J. L., Mah, S. M., McEwen, L. M., Saw, S.-M., Godfrey, K. M., Chong, Y.-S., Kwek, K., Kwoh, C.-K., Soh, S.-E., Chong, M. F. F., Barton, S., Karnani, N., Cheong, C. Y., … Holbrook, J. D. (2014). The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Research*, *24*(7), 1064–1074. https://doi.org/10.1101/gr.171439.113

Temple, J. L., Bernard, C., Lipshultz, S. E., Czachor, J. D., Westphal, J. A., & Mestre, M. A. (2017). The Safety of Ingested Caffeine: A Comprehensive Review. *Frontiers in Psychiatry*, *8*. https://doi.org/10.3389/fpsyt.2017.00080

Teschendorff, A. E., Marabita, F., Lechner, M., Bartlett, T., Tegner, J., Gomez-Cabrero, D., & Beck, S. (2013). A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics (Oxford, England)*, *29*(2), 189–196. https://doi.org/10.1093/bioinformatics/bts680

Teschendorff, A. E., West, J., & Beck, S. (2013). Age-associated epigenetic drift: Implications, and a case of epigenetic thrift? *Human Molecular Genetics*, *22*(R1), R7–R15. https://doi.org/10.1093/hmg/ddt375

Thapar, A., Rice, F., Hay, D., Boivin, J., Langley, K., van den Bree, M., Rutter, M., & Harold, G. (2009). Prenatal Smoking Might Not Cause Attention-Deficit/Hyperactivity Disorder: Evidence from a Novel Design. *Biological Psychiatry*, *66*(8), 722–727. https://doi.org/10.1016/j.biopsych.2009.05.032

Thapar, A., & Rutter, M. (2009). Do prenatal risk factors cause psychiatric disorder? Be wary of causal claims. *British Journal of Psychiatry*, *195*(2), 100–101. https://doi.org/10.1192/bjp.bp.109.062828

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, E. M., Esko, T., Nalls, M. A., Ganna, A., Paynter, N., Monda, K. L., Amin, N., Fischer, K., Renstrom, F., Ngwa, J. S., Huikari, V., Cavadino, A., Nolte, I. M., Teumer, A., … Chasman, D. I. (2015). Genome-wide meta-analysis identifies six novel loci associated with

habitual coffee consumption. *Molecular Psychiatry*, *20*(5), 647–656. https://doi.org/10.1038/mp.2014.107

Thompson, F. E., & Subar, A. F. (2017). Dietary Assessment Methodology. In *Nutrition in the Prevention and Treatment of Disease* (pp. 5–48). Elsevier. https://doi.org/10.1016/B978-0-12-802928-2.00001-1

Tiesler, C. M. T., & Heinrich, J. (2014). Prenatal nicotine exposure and child behavioural problems. *European Child & Adolescent Psychiatry*, *23*(10), 913–929. https://doi.org/10.1007/s00787-014-0615-y

Tobacco Free Initiative (World Health Organization). (2013). *WHO recommendations for the prevention and management of tobacco use and second-hand smoke exposure in pregnancy*. http://www.ncbi.nlm.nih.gov/books/NBK190304/

Treur, J. L., Demontis, D., Smith, G. D., Sallis, H., Richardson, T. G., Wiers, R. W., Børglum, A. D., Verweij, K. J. H., & Munafò, M. R. (2019). Investigating causality between liability to ADHD and substance use, and liability to substance use and ADHD risk, using Mendelian randomization. *Addiction Biology*, e12849. https://doi.org/10.1111/adb.12849

Treur, J. L., Taylor, A. E., Ware, J. J., McMahon, G., Hottenga, J.-J., Baselmans, B. M. L., Willemsen, G., Boomsma, D. I., Munafò, M. R., & Vink, J. M. (2016). Associations between smoking and caffeine consumption in two European cohorts: Smoking and caffeine consumption. *Addiction*, *111*(6), 1059–1068. https://doi.org/10.1111/add.13298

Treur, J. L., Taylor, A. E., Ware, J. J., Nivard, M. G., Neale, M. C., McMahon, G., Hottenga, J.-J., Baselmans, B. M. L., Boomsma, D. I., Munafò, M. R., & Vink, J. M. (2017). Smoking and caffeine consumption: A genetic analysis of their association: Smoking and caffeine. *Addiction Biology*, *22*(4), 1090–1102. https://doi.org/10.1111/adb.12391

Tukey, J. W. (1977). *Exploratory data analysis* (Ser. Addison-wesley series in behavioral science). Addison-Wesley Pub.

Turecki, G., & Meaney, M. J. (2016). Effects of the Social Environment and Stress on Glucocorticoid Receptor Gene Methylation: A Systematic Review. *Biological Psychiatry*, *79*(2), 87–96. https://doi.org/10.1016/j.biopsych.2014.11.022

Twenge, J. M., Cooper, A. B., Joiner, T. E., Duffy, M. E., & Binau, S. G. (2019).

Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. *Journal of Abnormal Psychology*, *128*(3), 185–199. https://doi.org/10.1037/abn0000410

Tyrrell, J., Huikari, V., Christie, J. T., Cavadino, A., Bakker, R., Brion, M.-J. A., Geller, F., Paternoster, L., Myhre, R., Potter, C., Johnson, P. C. D., Ebrahim, S., Feenstra, B., Hartikainen, A.-L., Hattersley, A. T., Hofman, A., Kaakinen, M., Lowe, L. P., Magnus, P., … Freathy, R. M. (2012). Genetic variation in the 15q25 nicotinic acetylcholine receptor gene cluster (CHRNA5–CHRNA3–CHRNB4) interacts with maternal self-reported smoking status during pregnancy to influence birth weight. *Human Molecular Genetics*, *21*(24), 5344–5358. https://doi.org/10.1093/hmg/dds372

van Dam, R. M., Hu, F. B., & Willett, W. C. (2020). Coffee, Caffeine, and Health. *New England Journal of Medicine*, *383*(4), 369–378. https://doi.org/10.1056/NEJMra1816604

van Dongen, J., Hagenbeek, F. A., Suderman, M., Roetman, P. J., Sugden, K., Chiocchetti, A. G., Ismail, K., Mulder, R. H., Hafferty, J. D., Adams, M. J., Walker, R. M., Morris, S. W., Lahti, J., Küpers, L. K., Escaramis, G., Alemany, S., Jan Bonder, M., Meijer, M., Ip, H. F., … Boomsma, D. I. (2021). DNA methylation signatures of aggression and closely related constructs: A meta-analysis of epigenome-wide studies across the lifespan. *Molecular Psychiatry*, *26*(6), 1–15. https://doi.org/10.1038/s41380-020-00987-x

van der Knaap, L. J., Riese, H., Hudziak, J. J., Verbiest, M. M. P. J., Verhulst, F. C., Oldehinkel, A. J., & van Oort, F. V. A. (2014). Glucocorticoid receptor gene (NR3C1) methylation following stressful events between birth and adolescence. The TRAILS study. *Translational Psychiatry*, *4*(4), e381–e381. https://doi.org/10.1038/tp.2014.22

Van der Most, P. J., Küpers, L. K., Snieder, H., & Nolte, I. (2017). QCEWAS: Automated quality control of results of epigenome-wide association studies. *Bioinformatics*, *33*(8), 1243–1245. https://doi.org/10.1093/bioinformatics/btw766

van der Sijde, M. R., Ng, A., & Fu, J. (2014). Systems genetics: From GWAS to

disease pathways. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1842*(10), 1903–1909. https://doi.org/10.1016/j.bbadis.2014.04.025

van Dongen, J., Nivard, M. G., Willemsen, G., Hottenga, J.-J., Helmer, Q., Dolan, C. V., Ehli, E. A., Davies, G. E., van Iterson, M., Breeze, C. E., Beck, S., Suchiman, H. E., Jansen, R., Meurs, J. B. van, Heijmans, B. T., Slagboom, P. E., & Boomsma, D. I. (2016). Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nature Communications*, *7*(1), 11115. https://doi.org/10.1038/ncomms11115

van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D. A., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child & Adolescent Psychiatry*, *12*(6), 281–289. https://doi.org/10.1007/s00787-003-0341-3

Vehmeijer, F. O. L., Küpers, L. K., Sharp, G. C., Salas, L. A., Lent, S., Jima, D. D., Tindula, G., Reese, S., Qi, C., Gruzieva, O., Page, C., Rezwan, F. I., Melton, P. E., Nohr, E., Escaramís, G., Rzehak, P., Heiskala, A., Gong, T., Tuominen, S. T., … Felix, J. F. (2020). DNA methylation and body mass index from birth to adolescence: Meta-analyses of epigenome-wide association studies. *Genome Medicine*, *12*(1), 105. https://doi.org/10.1186/s13073-020-00810-w

Vermeulen, J. M., Wootton, R. E., Treur, J. L., Sallis, H. M., Jones, H. J., Zammit, S., van den Brink, W., Goodwin, G. M., de Haan, L., & Munafò, M. R. (2019). Smoking and the risk for bipolar disorder: Evidence from a bidirectional Mendelian randomisation study. *The British Journal of Psychiatry*, 1–7. https://doi.org/10.1192/bjp.2019.202

Verster, J. C., & Koenig, J. (2018). Caffeine intake and its sources: A review of national representative studies. *Critical Reviews in Food Science and Nutrition*, *58*(8), 1250–1259. https://doi.org/10.1080/10408398.2016.1247252

Verweij, K. J. H., Treur, J. L., & Vink, J. M. (2018). Investigating causal associations between use of nicotine, alcohol, caffeine and cannabis: A two-sample bidirectional Mendelian randomization study. *Addiction*, *113*(7), 1333–1338. https://doi.org/10.1111/add.14154

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor

package. *Journal of Statistical Software*, *36*(3), 1–48.

Vineis, P., & Perera, F. (2007). Molecular Epidemiology and Biomarkers in Etiologic Cancer Research: The New in Light of the Old. *Cancer Epidemiology Biomarkers & Prevention*, *16*(10), 1954–1965. https://doi.org/10.1158/1055-9965.EPI-07-0457

Vineis, P., Veldhoven, K. van, Chadeau-Hyam, M., & Athersuch, T. J. (2013). Advancing the application of omics-based biomarkers in environmental epidemiology. *Environmental and Molecular Mutagenesis*, *54*(7), 461–467. https://doi.org/10.1002/em.21764

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *American Journal of Human Genetics*, *101*(1), 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Wadhwa, P. D., Buss, C., Entringer, S., & Swanson, J. M. (2009). Developmental Origins of Health and Disease: Brief History of the Approach and Current Focus on Epigenetic Mechanisms. *Seminars in Reproductive Medicine*, *27*(5), 358–368. https://doi.org/10.1055/s-0029-1237424

Wald, A. (1940). The Fitting of Straight Lines if Both Variables are Subject to Error. *The Annals of Mathematical Statistics*, *11*(3), 284–300.

Walton, E., Calhoun, V., Heijmans, B., Thompson, P., & Cecil, C. (2020). *The rise of neuroimaging epigenetics: A systematic review of studies examining associations between DNA methylation and brain imaging*.

Walton, E., Hass, J., Liu, J., Roffman, J. L., Bernardoni, F., Roessner, V., Kirsch, M., Schackert, G., Calhoun, V., & Ehrlich, S. (2016). Correspondence of DNA Methylation Between Blood and Brain Tissue and Its Application to Schizophrenia Research. *Schizophrenia Bulletin*, *42*(2), 406–414. https://doi.org/10.1093/schbul/sbv074

Walton, E., Pingault, J.-B., Cecil, C. A., Gaunt, T. R., Relton, C., Mill, J., & Barker, E. D. (2017). Epigenetic profiling of ADHD symptoms trajectories: A prospective, methylome-wide study. *Molecular Psychiatry*, *22*(2), 250–256. https://doi.org/10.1038/mp.2016.85

Walton, E., Relton, C. L., & Caramaschi, D. (2019a). Using Openly Accessible Resources to Strengthen Causal Inference in Epigenetic Epidemiology of Neurodevelopment and Mental Health. *Genes*, *10*(3), 193.

https://doi.org/10.3390/genes10030193

Ware, J. J., Tanner, J., Taylor, A. E., Bin, Z., Haycock, P., Bowden, J., Rogers, P. J., Davey Smith, G., Tyndale, R. F., & Munafò, M. R. (2017). Does coffee consumption impact on heaviness of smoking? *Addiction (Abingdon, England)*, *112*(10), 1842–1853. https://doi.org/10.1111/add.13888

Waszczuk, M. A., Zavos, H. M. S., Gregory, A. M., & Eley, T. C. (2014). The Phenotypic and Genetic Structure of Depression and Anxiety Disorder Symptoms in Childhood, Adolescence, and Young Adulthood. *JAMA Psychiatry*, *71*(8), 905. https://doi.org/10.1001/jamapsychiatry.2014.655

Weaver, I. C. G., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., Dymov, S., Szyf, M., & Meaney, M. J. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, *7*(8), 847–854. https://doi.org/10.1038/nn1276

Webb, P., Bain, C., & Page, A. (2017). *Essential Epidemiology: An Introduction for Students and Health Professionals*. Cambridge University Press.

Weder, N., Zhang, H., Jensen, K., Yang, B. Z., Simen, A., Jackowski, A., Lipschitz, D., Douglas-Palumberi, H., Ge, M., Perepletchikova, F., O'Loughlin, K., Hudziak, J. J., Gelernter, J., & Kaufman, J. (2014). Child Abuse, Depression, and Methylation in Genes Involved With Stress, Neural Plasticity, and Brain Circuitry. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(4), 417-424.e5. https://doi.org/10.1016/j.jaac.2013.12.025

Weinhold, L., Wahl, S., Pechlivanis, S., Hoffmann, P., & Schmid, M. (2016). A statistical model for the analysis of beta values in DNA methylation studies. *BMC Bioinformatics*, *17*. https://doi.org/10.1186/s12859-016-1347-4

Weissman, M. M., Wolk, S., Wickramaratne, P., Goldstein, R. B., Adams, P., Greenwald, S., Ryan, N. D., Dahl, R. E., & Steinberg, D. (1999). Children with prepubertal-onset major depressive disorder and anxiety grown up. *Archives of General Psychiatry*, *56*(9), 794–801.

Weng, X., Odouli, R., & Li, D.-K. (2008). Maternal caffeine consumption during pregnancy and the risk of miscarriage: A prospective cohort study. *American Journal of Obstetrics and Gynecology*, *198*(3), 279.e1-279.e8. https://doi.org/10.1016/j.ajog.2007.10.803

Wickström, R. (2007). Effects of Nicotine During Pregnancy: Human and Experimental Evidence. *Current Neuropharmacology*, *5*(3), 213–222. https://doi.org/10.2174/157015907781695955

Wigginton, B., & Lee, C. (2013). Stigma and hostility towards pregnant smokers: Does individuating information reduce the effect? *Psychology & Health*, *28*(8), 862–873. https://doi.org/10.1080/08870446.2012.762101

Wiklund, P., Karhunen, V., Richmond, R. C., Parmar, P., Rodriguez, A., De Silva, M., Wielscher, M., Rezwan, F. I., Richardson, T. G., Veijola, J., Herzig, K.-H., Holloway, J. W., Relton, C. L., Sebert, S., & Järvelin, M.-R. (2019). DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. *Clinical Epigenetics*, *11*(1), 97. https://doi.org/10.1186/s13148-019-0683-4

Wilkinson, P. (2009). Conceptualization about internalizing problems in children and adolescents. *Ciência & Saúde Coletiva*, *14*(2), 373–381. https://doi.org/10.1590/S1413-81232009000200007

Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, *26*(17), 2190–2191. https://doi.org/10.1093/bioinformatics/btq340

Wolke, D., Waylen, A., Samara, M., Steer, C., Goodman, R., Ford, T., & Lamberts, K. (2009). Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *British Journal of Psychiatry*, *195*(3), 249–256. https://doi.org/10.1192/bjp.bp.108.053751

Wootton, R. E., Richmond, R. C., Stuijfzand, B. G., Lawn, R. B., Sallis, H. M., Taylor, G. M. J., Hemani, G., Jones, H. J., Zammit, S., Davey Smith, G., & Munafò, M. R. (2019). Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: A Mendelian randomisation study. *Psychological Medicine*, 1–9. https://doi.org/10.1017/S0033291719002678

World Health Organization (Ed.). (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. World Health Organization.

Wray, N. R., Lin, T., Austin, J., McGrath, J. J., Hickie, I. B., Murray, G. K., & Visscher, P. M. (2021). From Basic Science to Clinical Application of Polygenic Risk Scores: A Primer. *JAMA Psychiatry*, *78*(1), 101–109.

https://doi.org/10.1001/jamapsychiatry.2020.3049

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S.-A., Bækvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttenschøn, H. N., Bybjerg-Grauholm, J., … Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*(5), 668–681. https://doi.org/10.1038/s41588-018-0090-3

Wright, J., Small, N., Raynor, P., Tuffnell, D., Bhopal, R., Cameron, N., Fairley, L., Lawlor, D. A., Parslow, R., Petherick, E. S., Pickett, K. E., Waiblinger, D., West, J., & on behalf of the Born in Bradford Scientific Collaborators Group. (2013). Cohort Profile: The Born in Bradford multi-ethnic family cohort study. *International Journal of Epidemiology*, *42*(4), 978–991. https://doi.org/10.1093/ije/dys112

Wu, D.-M., He, Z., Ma, L.-P., Wang, L.-L., Ping, J., & Wang, H. (2015). 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. *Toxicology and Applied Pharmacology*, *285*(2), 89–97. https://doi.org/10.1016/j.taap.2015.03.028

Xu, C.-J., Bonder, M. J., Söderhäll, C., Bustamante, M., Baïz, N., Gehring, U., Jankipersadsing, S. A., van der Vlies, P., van Diemen, C. C., van Rijkom, B., Just, J., Kull, I., Kere, J., Antó, J. M., Bousquet, J., Zhernakova, A., Wijmenga, C., Annesi-Maesano, I., Sunyer, J., … Koppelman, G. H. (2017). The emerging landscape of dynamic DNA methylation in early childhood. *BMC Genomics*, *18*(1), 25. https://doi.org/10.1186/s12864-016-3452-1

Xu, D., Zhang, B., Liang, G., Ping, J., Kou, H., Li, X., Xiong, J., Hu, D., Chen, L., Magdalou, J., & Wang, H. (2012). Caffeine-Induced Activated Glucocorticoid Metabolism in the Hippocampus Causes Hypothalamic-Pituitary-Adrenal Axis Inhibition in Fetal Rats. *PLoS ONE*, *7*(9), e44497. https://doi.org/10.1371/journal.pone.0044497

Xu, Z., Seidler, F. J., Ali, S. F., Slikker, W., & Slotkin, T. A. (2001). Fetal and

adolescent nicotine administration: Effects on CNS serotonergic systems. *Brain Research*, *914*(1), 166–178. https://doi.org/10.1016/S0006-8993(01)02797-4

Yanes, T., Young, M.-A., Meiser, B., & James, P. A. (2020). Clinical applications of polygenic breast cancer risk: A critical review and perspectives of an emerging field. *Breast Cancer Research*, *22*(1), 21. https://doi.org/10.1186/s13058-020-01260-3

Yang, A., Palmer, A. A., & de Wit, H. (2010). Genetics of caffeine consumption and responses to caffeine. *Psychopharmacology*, *211*(3), 245–257. https://doi.org/10.1007/s00213-010-1900-1

Yang, J., Weedon, M. N., Purcell, S., Lettre, G., Estrada, K., Willer, C. J., Smith, A. V., Ingelsson, E., O'Connell, J. R., Mangino, M., Mägi, R., Madden, P. A., Heath, A. C., Nyholt, D. R., Martin, N. G., Montgomery, G. W., Frayling, T. M., Hirschhorn, J. N., McCarthy, M. I., … Visscher, P. M. (2011). Genomic inflation factors under polygenic inheritance. *European Journal of Human Genetics*, *19*(7), 807–812. https://doi.org/10.1038/ejhg.2011.39

York, T. P., Jackson-Cook, C., Moyer, S., Roberson-Nay, R., Murphy, S. K., Fuemmeler, B. F., Latendresse, S. J., Lapato, D. M., Wolen, A. R., Do, E. K., Hoyo, C., & Strauss, J. F. (2019). Replicated Umbilical Cord Blood DNA Methylation Loci Associated with Gestational Age at Birth. *BioRxiv*, 749135. https://doi.org/10.1101/749135

Young, S. E., Stallings, M. C., Corley, R. P., Krauter, K. S., & Hewitt, J. K. (2000). Genetic and environmental influences on behavioral disinhibition. *American Journal of Medical Genetics*, 684–695.

Yousefi, P., Huen, K., Davé, V., Barcellos, L., Eskenazi, B., & Holland, N. (2015). Sex differences in DNA methylation assessed by 450 K BeadChip in newborns. *BMC Genomics*, *16*. https://doi.org/10.1186/s12864-015-2034-y

Yousefi, P., Huen, K., Quach, H., Motwani, G., Hubbard, A., Eskenazi, B., & Holland, N. (2015). Estimation of blood cellular heterogeneity in newborns and children for epigenome-wide association studies. *Environmental and Molecular Mutagenesis*, *56*(9), 751–758. https://doi.org/10.1002/em.21966

Yu, T., Campbell, S. C., Stockmann, C., Tak, C., Schoen, K., Clark, E. A. S., Varner, M. W., Spigarelli, M. G., & Sherwin, C. M. T. (2016). Pregnancy-induced changes in the pharmacokinetics of caffeine and its metabolites. *The Journal of Clinical Pharmacology*, *56*(5), 590–596. https://doi.org/10.1002/jcph.632

Zhu, S.-H., & Valbø, A. (2002). Depression and smoking during pregnancy. *Addictive Behaviors*, *27*(4), 649–658. https://doi.org/10.1016/S0306-4603(01)00199-X

Zuccolo, L., Lewis, S. J., Davey Smith, G., Sayal, K., Draper, E. S., Fraser, R., Barrow, M., Alati, R., Ring, S., Macleod, J., Golding, J., Heron, J., & Gray, R. (2013). Prenatal alcohol exposure and offspring cognition and school performance. A 'Mendelian randomization' natural experiment. *International Journal of Epidemiology*, *42*(5), 1358–1370. https://doi.org/10.1093/ije/dyt172

Appendices

Appendix A

A1 List of phenotypes included in the study.

Phenotype	Assessment instrument	Time point
Offspring: Children		
Mental health		
ADHD symptoms (categorical)	SDQ[1]	6.7 years
Conduct disorders symptoms (categorical)	SDQ	6.7 years
Oppositional-defiant disorder symptoms (categorical)	DAWBA[2]	7.5 years
SMFQ score (categorical)	SMFQ[3]	9.5 years
SDQ emotional symptoms score (categorical)	SDQ	6.8 years
General anxiety symptoms score (categorical)	DAWBA	7.5 years
Total behavioural difficulties score (categorical)	SDQ	6.7 years
Specific phobia clinical diagnosis (binary)	DAWBA	9.5 years
Autism diagnosis (binary)	Derived combining multiple measures[4]	9 years
Non-mental health		
Sleep duration in hours	Maternal report	6.7 years
Number of life events (categorical)	Life events inventory (maternal report)	6.7 years
Problems with sleep initiation in past year (binary)	Maternal report	6.7 years
Problems with sleep maintenance in past year (binary)	Maternal report	6.7 years
IQ total score	WISC[5]	8 years
Body mass Index		7 years
Right or left-handed (binary)	Child completed	11 years
Caffeine		
Total mg/day caffeine from tea, cola, coffee	Maternal report	8 years
Child drinks caffeinated tea (binary)	Maternal report	8 years
Child drinks caffeinated coffee (binary)	Maternal report	8 years
Offspring: Adolescents		
Mental health		
ADHD symptoms (categorical)	SDQ	16.6 years
Conduct disorder symptoms (categorical)	SDQ	16.6 years
Oppositional-defiant disorder symptoms (categorical)	DAWBA	15.5 years
Depressive symptoms – Sum of all the 5 depression symptoms (categorical)	CIS-R[6]	18 years
Depression symptoms total score (categorical)	MFQ[7]	17.5 years
Depression symptoms score (categorical)	MFQ	14 years

[1] The Strengths and Difficulties Questionnaire
[2] The Development and Well-Being Assessment
[3] The Short Mood and Feelings Questionnaire
[4] Colin, D., Jean, G., & Patrick, F. (2010). Traits contributing to the autistic spectrum. *Plos One, 5*(9). doi:10.1371/journal.pone.0012633
[5] The Wechsler-Intelligence Scale
[6] The Revised Clinical Interview Schedule
[7] The Mood and Feelings Questionnaire

PTSD (self-report 6-band computer prediction, binary)	DAWBA	15 years
Self-harming behaviour with suicidal intent (binary)	Derived from multiple measures[8]	15 years
Emotional symptoms score (categorical)	SDQ	16.5 years
Anxiety score (categorical)	CIS-R	17.1 years
Phobias: Phobia symptom score (categorical)	CIS-R	17.1 years
Total behavioural difficulties score (categorical)	SDQ	16.5 years
Ever treated for an eating disorder (binary)	Child completed	13 years
Psychosis positive symptoms (categorical)	Psychosis interview	12 years
Psychosis negative symptoms score	PLIKS[9]	16.5 years
Psychosis positive symptoms (categorical)	Psychosis interview	18 years
Ever treated for an eating disorder (binary)	Child completed	16 years
Non-mental health		
Big-5 personality traits: Extraversion	IPIP[10]	13 years
Big-5 personality traits: Agreeableness	IPIP	13 years
Big-5 personality traits: Conscientiousness	IPIP	13 years
Big-5 personality traits: Emotional Stability (neuroticism)	IPIP	13 years
Big-5 personality traits: Intellect	IPIP	13 years
Maintaining sleep: Number of times young person usually wakes up at night (categorical)	Child completed	15 years
Initiating sleep: Average time (minutes) YP takes to fall asleep per week (categorical)	Child completed	15 years
Frequency respondent did any exercise during the past year (categorical)	Child completed	14 years
GCSE grades A-C (binary)	Child completed	18 years
GCSE grades D-G (binary)	Child completed	18 years
BMI		17 years
IQ total score	WASI[11]	15.5 years
Sleep duration	Child completed	15.5 years
Number of life events (categorical)	Life events inventory (child self-report)	16.5 years
Substance use		
Alcohol		
AUDIT: Frequency young person has a drink containing alcohol (Continuous)	Child Self-report	17.1 years
Level of risk identified by alcohol use disorders identification test	AUDIT[12]	17.1 years
Frequency had 6+ drinks on one occasion	Child Self-report	17.1 years
No. full drinks needed to feel tipsy/have buzz over last 3 months (categorical)	Child Self-report	17.1 years
Number of alcoholic drinks on a typical day (categorical)	Child Self-report	18 years
Total score of AUDIT test (categorical)	AUDIT	18 years
Number of times had whole drink in the past 6 months (categorical)	Self-report	12 years
Number of drinks took to feel different after first 5 times drinking (categorical)	Self-report	12 years
Number of times had 3+ drinks in one day (categorical)	Self-report	12 years
Tobacco		

[8] Easey, K.E., Mars, B., Pearson, R. et al. Eur Child Adolesc Psychiatry (2019) 28: 1079. https://doi.org/10.1007/s00787-018-1266-1
[9] The psychosis-like symptoms measure
[10] The International Personality Item Pool
[11] Wechsler abbreviated scale of Intelligence
[12] Alcohol Use Disorders Identification Test

Age of respondent when first smoked a cigarette	Self-report	14 years
Age when respondent smoked first whole cigarette (years)	Self-report	18 years
Number of cigarettes respondent smoked altogether in lifetime (categorical)	Self-report	18 years
Frequency young person smokes cannabis (categorical)	Self-report	16.5 years
Respondent has smoked a cigarette (including roll-ups) (binary)	Self-report	14 years
Total number of cigarettes that the respondent has smoked (binary)	Self-report	14 years
Respondent has ever smoked a whole cigarette (including roll-ups) (binary)	Self-report	18 years
Young person has ever tried cannabis (binary)	Self-report	16.5 years
Caffeine		
Total mg/day caffeine from tea, coffee, cola (categorical)	Maternal report	13 years
Tea mg/day caffeine teenager	Maternal report	13 years
Coffee mg/day caffeine teenager (categorical)	Maternal report	13 years
Cola mg/day caffeine teenager (categorical)	Maternal report	13 years
Mothers during pregnancy		
Mental health		
Depression symptoms (binary)	EPDS[13]	18 weeks gest
Depression symptoms (binary)	EPDS	32 weeks gest
Hypersensitivity to interpersonal rejection	IPMS[14]	18 weeks gest
Anxiety symptoms (binary)	CCEI[15]	
Non-mental health		
Number of life events mother experienced in pregnancy (categorical)	Life events Inventory	18 weeks gest
Image perception score during pregnancy	Self-report	18 weeks gest
Image perception change from before to during pregnancy	Self-report	18 weeks gest
Your reactions to becoming a parent (categorical)	Self-report	18 weeks gest
Activity level compared with other pregnant women (categorical)	Self-report	32 weeks gest
Physical activity (binary)	Self-report	32 weeks gest
Vomited in first three months of pregnancy (binary)	Self-report	18 weeks gest
Social class based on occupation (categorical)	Self-report	32 weeks gest
Mothers' highest education in pregnancy (categorical)	Self-report	32 weeks gest
Substance use		
Alcohol		
Alcohol: binging (categorical)	Self-report	18 weeks gest
Alcohol per week	Self-report	32 weeks gest
Alcohol: binging (categorical)	Self-report	32 weeks gest
Tobacco		
Smoking first three months in pregnancy (binary)	Self-report	18 weeks gest
Ever smoked during pregnancy (binary)	Self-report	8 weeks gest
Stopped smoking during pregnancy (binary)	Self-report	8 weeks gest
Cut down smoking during pregnancy (binary)	Self-report	8 weeks gest
Caffeine		
Total mg/day caffeine pregnancy	Self-report	18 weeks gest
Tea mg/day caffeine pregnancy (categorical)	Self-report	18 weeks gest
Coffee mg/day caffeine pregnancy (categorical)	Self-report	18 weeks gest

[13] Edinburgh Postnatal Depression Scale
[14] Interpersonal Sensitivity Measure
[15] The Crown Crisp Experiential Index (anxiety sub-scale)

Cola mg/day caffeine pregnancy (categorical)	Self-report	18 weeks gest
Total mg/day caffeine pregnancy	Self-report	32 weeks gest
Tea mg/day caffeine pregnancy	Self-report	32 weeks gest
Coffee mg/day caffeine pregnancy (categorical)	Self-report	32 weeks gest
Cola mg/day caffeine pregnancy (categorical)	Self-report	32 weeks gest
Consumed more caffeine during pregnancy (binary)	Self-report	8 weeks gest
Never has been drinking caffeine (binary)	Self-report	8 weeks gest
Did not change caffeine consumption during pregnancy (binary)	Self-report	8 weeks gest
Reduced caffeine consumption during pregnancy (binary)	Self-report	8 weeks gest
Never drank tea vs. drinking tea (binary)	Self-report	8 weeks gest
Stopped drinking tea during pregnancy (binary)	Self-report	8 weeks gest
Reduced tea consumption during pregnancy (binary)	Self-report	8 weeks gest
Craved or had more tea during pregnancy (binary)	Self-report	8 weeks gest
Never drank coffee vs. drinking coffee (binary)	Self-report	8 weeks gest
Stopped drinking coffee during pregnancy (binary)	Self-report	8 weeks gest
Reduced coffee consumption during pregnancy	Self-report	8 weeks gest
Craved or had more coffee during pregnancy (binary)	Self-report	8 weeks gest
Never drank cola vs. drinking cola (binary)	Self-report	8 weeks gest
Stopped drinking cola during pregnancy (binary)	Self-report	8 weeks gest
Reduced cola consumption during pregnancy (binary)	Self-report	8 weeks gest
Craved or had more cola during pregnancy (binary)	Self-report	8 weeks gest
Cutting down cola consumption during pregnancy (binary)	Self-report	8 weeks gest
Other substances		
Hard drugs (binary)	Self-report	18 weeks gest
Cannabis first three months in pregnancy (binary)	Self-report	8 weeks gest
Mothers outside of pregnancy		
Mental health		
Anxiety symptoms (binary)	CCEI	11 years (child age)
Depression symptoms (binary)	EPDS	11 years
Ever had bulimia (binary)	Self-report	12 weeks gest
Ever had drug addiction (binary)	Self-report	12 weeks gest
Ever had alcoholism (binary)	Self-report	12 weeks gest
Ever had schizophrenia (binary)	Self-report	12 weeks gest
Ever had anorexia nervosa (binary)	Self-report	12 weeks gest
Ever had severe depression (binary)	Self-report	12 weeks gest
Ever had other psychiatric problem (binary)	Self-report	12 weeks gest
Image perception 3 months before pregnancy (categorical)	Self-report	18 weeks gest
Non-mental health		
Number of life events mums (categorical)	Life-events inventory (self-report)	11 years
Impulsivity trait	KSP[16]	9 years
Monotony avoidance trait	KSP	9 years
Anger trait	KSP	9 years
Suspicion trait	KSP	9 years
Detachment trait	KSP	9 years
Social class based on occupation (categorical)	Self-report	4 years

[16] The Karolinska Scale of Personality

Mothers' highest educational qualifications (categorical)	Self-report	5 years
BMI mothers	Self-report	12 weeks gest
Mother participates in physical activity (binary)	Self-report	18 years
Substance use		
Tobacco		
Mother has ever been smoker (binary)	Self-report	18 weeks gest
Number of cigarettes mother smoked before pregnancy	Self-report	18 weeks gest
Number of cigarettes mother has smoked last 2 weeks	Self-report	8 years
Caffeine		
Daily caffeine intake from cola (mg) (categorical)	Self-report	8 years
Daily caffeine intake from tea (mg) (categorical)	Self-report	8 years
Daily caffeine intake from coffee (mg) (categorical)	Self-report	8 years
Mothers' daily caffeine intake through tea, coffee & cola (including persons with missing 1 or 2 drinks)	Self-report	8 years
Alcohol		
Mothers' total alcohol units daily (categorical)	Self-report	8 years
Mothers' pre-pregnancy drinking (never/ever) (binary)	Self-report	18 weeks gest
Number of days in past month that mother had at least 4 units of alcohol (categorical)	Self-report	5 years
Mothers' total alcohol units daily (categorical)	Self-report	4 years
AUDIT score in mothers (based on risk level) (categorical)	AUDIT	18 years

Appendix B

B1 Two-sample t-tests (two-sided) to test differences between participants with complete and partially missing genotype data.

	Participants with genetic data*	**N**	**Mean (SE)**	**P-value diff**
Caffeine consumption during pregnancy (mg/day)	Mothers and offspring	4918	160 (1.76)	0.0002
	Mothers or offspring	4778	151 (1.60)	
Social class*	Mothers and offspring	4052	1.98 (0.02)	<0.001
	Mothers or offspring	3733	1.74 (0.02)	
Maternal education**	Mothers and offspring	3935	2.1 (0.02)	<0.001
	Mothers or offspring	3353	2.4 (0.02)	
Maternal age (years)	Mothers and offspring	4788	27.57 (0.07)	<0.001
	Mothers or offspring	4938	28.63 (0.07)	

*Note. Mothers and offspring refers to mother-offspring pairs that both have genotype data in ALSPAC. Mothers or offspring refers to mother-offspring pairs where either mother or offspring have genotype data but not both. * Social class levels are based on individual's occupation where classes I to V stands for occupations: I – professional; II – managerial and technical; III – skilled non-manual and manual; IV – partly-skilled; V – unskilled; ** 4 categories: 0 = none or CSE, 1 = vocational, 2 = O-level, 3 = A-level; 4 = degree (CSE reflects to the certificate of secondary education which is available for both academic and vocational subjects. O level is equivalent to grades D and E and A level is equivalent to grades A to C after GCSE (General Certificate of Secondary Education) examination. Degree level reflects to higher education diploma.*

B2 Chi-Square test to test differences between participants with complete and partially missing genotype data.

	Participants with genetic data		
Maternal smoking during the 1st trimester of pregnancy (yes/no)	**Mothers and offspring**	**Mothers or offspring**	**Total N**
Yes	N = 924	N = 1,304	2,228
No	N = 3,864	N = 3,634	7,498
Total N	4,788	4,938	9,726
Results: X2(1) = 69.57, P < 0.001			

Note. Mothers and offspring refers to mother-offspring pairs that both have genotype data in ALSPAC. Mothers or offspring refers to mother-offspring pairs where either mother or offspring have genotype data but not both.

Appendix C

C1 Associations between the lifetime smoking PRS and smoking phenotypes in mothers during and outside of pregnancy and adolescents

Phenotype	Effect estimate	Effect size*	95% CI	P-value	Sample size	Adj. R^2**
Mothers during pregnancy						
Tobacco smoked in 1st three months of pregnancy	OR	1.235	1.159, 1.315	$9.41x10^{-6}$	7237	0.04
Mother cut down tobacco consumption	OR	1.168	1.097, 1.244	<0.001	7269	0.02
Mother stopped smoking during pregnancy	OR	0.871	0.775, 0.979	0.024	1863	0.01
Mothers outside of pregnancy						
Mother has ever smoked	OR	1.147	1.089, 1.209	<0.001	7194	0.01
Number of cigarettes mother smoked before pregnancy	Beta	0.194	0.124, 0.264	$5.27x10^{-8}$	3426	0.05
Offspring: Adolescents						
Smoked age 14 years	OR	1.117	1.033, 1.208	0.009	4145	0.03
Smoked more than 20 cigarettes age 14	OR	1.156	0.995, 1.342	0.057	1058	0.01
Age 1st smoked a cigarette (asked age 14)	Beta	-0.052	-0.096, -0.009	0.019	1064	0.01
Ever smoked a whole cigarette age 18	OR	1.130	1.035, 1.233	0.010	2402	0.01
Number of cigarettes smoked in lifetime age 18	Beta	0.084	-0.006, 0.174	0.069	1144	0.002

* Reflects the average change in the outcome that is associated with a one standard deviation increase in the PRS. For binary outcomes, this will be the odds ratio (e.g., Mother's odds of ever smoking are 1.147 times compared to not smoking), for continuous outcomes it represents the average unit change (e.g., 0.775 cigarettes smoked). ** For the logistic regression models pseudo R^2 is reported.

Appendix D

D1 Associations between the maternal and offspring smoking initiation PRS and offspring phenotypes <10 years.

		Intergenerational analyses						Childhood analyses					
		Regression analyses			Permutation testing			Regression analyses			Permutation testing		
Phenotype	**Effect estimate**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	**Sample size**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	**Sample size**
1. Total caffeine	Beta	0.045	0.021, 0.068	<0.001	<0.001, 0.004	<0.001	4067	0.032	0.010, 0.055	0.005	0.002, 0.013	0.006	4589
2. Anxiety	Beta	-0.033	-0.053, -0.012	0.002	<0.001, 0.007	0.002	4993	-0.031	-0.051, -0.010	0.003	<0.001, 0.007	0.002	5355
3. BMI	Beta	0.076	0.018, 0.135	0.010	0.007, 0.022	0.013	5032	0.050	<0.001, 0.101	0.051	0.036, 0.063	0.048	5799
4. IQ	Beta	-0.592	-1.049, -0.134	0.011	0.009, 0.026	0.016	4675	-0.735	-1.183, -0.287	0.001	<0.001, 0.004	<0.001	5295
5. Conduct disorder	Beta	0.024	0.004, 0.044	0.019	0.013, 0.032	0.021	5012	0.030	0.012, 0.049	0.001	<0.001, 0.006	0.001	5326
6. Handedness	OR	1.114	1.012, 1.225	0.030	0.006, 0.021	0.012	4849	1.045	0.954, 1.145	0.315	0.263, 0.320	0.291	5403
7. Specific phobia	OR	1.322	0.964, 1.813	0.078	0.042, 0.071	0.055	5100	1.182	0.881, 1.587	0.241	0.199, 0.252	0.225	5470
8. Emotional problems	Beta	-0.016	-0.037, 0.004	0.117	0.106, 0.148	0.126	5139	-0.011	-0.031, 0.009	0.267	0.236, 0.291	0.263	5459
9. ADHD	Beta	0.016	-0.013, 0.045	0.277	0.232, 0.287	0.259	4916	0.030	0.003, 0.058	0.030	0.024, 0.047	0.034	5219
10. Sleep duration	Beta	-0.009	-0.033, 0.014	0.426	0.392, 0.454	0.423	5127	-0.019	-0.042, 0.004	0.106	0.107, 0.149	0.127	5443
11. Behavioural difficulties	Beta	0.010	-0.021, 0.041	0.522	0.482, 0.544	0.513	5133	0.022	-0.008, 0.051	0.152	0.130, 0.176	0.152	5452
12. Depression	Beta	-0.006	-0.027, 0.015	0.557	0.524, 0.586	0.555	4885	-0.007	-0.027, 0.012	0.466	0.442, 0.504	0.473	5434
13. Sleep maintenance	OR	0.983	0.919, 1.051	0.589	0.534, 0.596	0.565	5127	0.973	0.913, 1.038	0.383	0.313, 0.372	0.342	5448
14. ODD	Beta	-0.004	-0.024, 0.016	0.700	0.683, 0.740	0.712	4943	0.015	-0.005, 0.034	0.148	0.146, 0.194	0.169	5319
15. Autism	OR	1.027	0.722, 1.460	0.874	0.860, 0.901	0.882	5975	1.153	0.803, 1.654	0.411	0.380, 0.442	0.411	6156
16. Sleep initiation	OR	0.995	0.934, 1.061	0.874	0.827, 0.873	0.851	5150	0.971	0.913, 1.032	0.309	0.269, 0.326	0.297	5476
17. Life events	Beta	<0.001	-0.018, 0.019	0.996	0.991, 0.999	0.997	5167	0.010	-0.008, 0.028	0.271	0.237, 0.292	0.264	5493

Note. The intergenerational analysis represents offspring phenotypes <10 years regressed on maternal PRS. The childhood analysis represents offspring phenotypes <10 years regressed on offspring PRS.

E1 Associations between maternal and offspring lifetime smoking PRS and offspring phenotypes <10 years.

		Intergenerational analyses						**Childhood analyses**					
		Regression analyses			**Permutation testing**			**Regression analyses**			**Permutation testing**		
Phenotype	**Effect estimate**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	**Sample size**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	**Sample size**
1. IQ	Beta	-0.742	-1.202, -0.282	0.002	<0.001, 0.007	0.002	4675	-0.929	-1.371, -0.488	$3.73\text{x}10^{-5}$	<0.001, 0.004	<0.001	5290
2. Conduct disorder	Beta	0.026	0.007, 0.045	0.009	0.003, 0.014	0.007	5012	0.029	0.010, 0.048	0.003	0.001, 0.009	0.003	5326
3. BMI	Beta	0.063	0.007, 0.119	0.029	0.020, 0.043	0.030	5032	0.026	-0.025, 0.076	0.316	0.282, 0.341	0.311	5799
4. Total caffeine	Beta	0.021	-0.003, 0.045	0.079	0.063, 0.097	0.079	4067	0.015	-0.007, 0.038	0.187	0.170, 0.220	0.194	4589
5. Sleep initiation	OR	0.950	0.892, 1.012	0.104	0.064, 0.099	0.080	5150	0.968	0.911, 1.029	0.273	0.203, 0.256	0.229	5476
6. Behavioural difficulties	Beta	0.025	-0.005, 0.056	0.107	0.089, 0.129	0.108	5133	0.045	0.016, 0.075	0.003	<0.001, 0.007	0.002	5452
7. ADHD	Beta	0.023	-0.006, 0.052	0.117	0.098, 0.139	0.117	4916	0.037	0.009, 0.065	0.009	0.006, 0.020	0.011	5219
8. Specific phobia	OR	1.222	0.916, 1.631	0.156	0.179, 0.230	0.204	5100	0.824	0.628, 1.083	0.150	0.169, 0.219	0.193	5470
9. Anxiety	Beta	-0.012	-0.033, 0.009	0.256	0.229, 0.284	0.256	4993	-0.014	-0.034, 0.007	0.189	0.150, 0.198	0.173	5355
10. Sleep duration	Beta	-0.013	-0.036, 0.010	0.259	0.243, 0.299	0.270	5127	0.002	-0.021, 0.024	0.878	0.851, 0.893	0.873	5443
11. Sleep maintenance	OR	1.019	0.952, 1.090	0.559	0.503, 0.565	0.534	5127	0.984	0.924, 1.048	0.594	0.556, 0.618	0.587	5448
12. Autism	OR	1.104	0.768, 1.589	0.563	0.512, 0.574	0.543	5975	1.259	0.838, 1.891	0.243	0.163, 0.213	0.187	6156
13. ODD	Beta	0.006	-0.014, 0.026	0.574	0.557, 0.619	0.588	4943	0.031	0.012, 0.051	0.002	<0.001, 0.004	<0.001	5319
14. Emotional problems	Beta	-0.005	-0.025, 0.016	0.656	0.630, 0.689	0.660	5139	-0.012	-0.031, 0.008	0.248	0.210, 0.264	0.236	5459
15. Depression	Beta	-0.003	-0.023, 0.018	0.809	0.783, 0.833	0.809	4885	0.010	-0.010, 0.030	0.323	0.300, 0.359	0.329	5434
16. Handedness	OR	1.009	0.914, 1.114	0.846	0.790, 0.839	0.815	4849	1.006	0.924, 1.096	0.876	0.866, 0.906	0.887	5399
17. Life events	Beta	-0.002	-0.020, 0.017	0.853	0.838, 0.882	0.861	5167	0.014	-0.004, 0.032	0.117	0.101, 0.143	0.121	5493

Note. Intergenerational analysis refers to maternal PRS predicting offspring phenotypes <10 years. Childhood analysis refers to offspring PRS predicting offspring phenotypes <10 years.

Appendix F

F1 Associations between maternal and offspring caffeine PRS and offspring phenotypes <10 years.

		Intergenerational analyses						Childhood analyses					
		Regression analyses			Permutation testing			Regression analyses			Permutation testing		
Phenotype	**Effect estimate**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	**Sample size**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	**Sample size**
1. Specific phobia	OR	0.724	0.519, 1.012	0.057	0.019, 0.040	0.028	5100	0.999	0.723, 1.381	0.997	0.993, 1.000	0.998	4900
2. Depression	Beta	-0.017	-0.039, 0.002	0.075	0.056, 0.089	0.071	4885	-0.017	-0.037, 0.002	0.081	0.055, 0.088	0.070	5434
3. ADHD	Beta	-0.018	-0.008, 0.050	0.161	0.110, 0.152	0.130	4916	-0.018	-0.046, 0.010	0.206	0.199, 0.251	0.224	5219
4. Handedness	OR	1.064	0.968, 1.169	0.178	0.143, 0.189	0.165	4849	0.980	0.897, 1.070	0.624	0.548, 0.610	0.579	5399
5. Life events	Beta	-0.008	-0.007, 0.030	0.228	0.217, 0.271	0.243	5167	-0.008	-0.027, 0.010	0.366	0.329, 0.390	0.359	5493
6. ODD	Beta	0.002	-0.032, 0.008	0.240	0.257, 0.314	0.285	4943	0.002	-0.018, 0.022	0.829	0.583, 0.644	0.614	5319
7. Sleep initiation	OR	0.972	0.913, 1.036	0.352	0.316, 0.375	0.345	5150	0.951	0.895, 1.010	0.094	0.059, 0.092	0.074	5476
8. Total caffeine	Beta	0.008	-0.015, 0.032	0.490	0.444, 0.506	0.475	4067	0.010	-0.012, 0.032	0.377	0.349, 0.410	0.379	4589
9. BMI	Beta	0.025	-0.033, 0.084	0.387	0.364, 0.425	0.394	5032	0.025	-0.027, 0.077	0.348	0.297, 0.356	0.326	5799
10. Behavioural difficulties	Beta	-0.015	-0.019, 0.043	0.441	0.369, 0.431	0.400	5133	-0.015	-0.044, 0.015	0.324	0.294, 0.353	0.323	5452
11. Emotional problems	Beta	0.001	-0.014, 0.027	0.538	0.569, 0.631	0.600	5139	0.001	-0.018, 0.021	0.883	0.881, 0.919	0.901	5459
12. Sleep duration	Beta	-0.026	-0.030, 0.017	0.577	0.544, 0.606	0.575	5127	-0.026	-0.048, -0.004	0.018	0.011, 0.028	0.018	5443
13. CD	Beta	-0.006	-0.024, 0.015	0.624	0.627, 0.686	0.657	5012	-0.006	-0.024, 0.013	0.563	0.541, 0.603	0.572	5326
14. Autism	OR	1.052	0.758, 1.461	0.742	0.733, 0.787	0.761	5975	0.850	0.603, 1.199	0.326	0.307, 0.366	0.336	6156
15. IQ	Beta	0.276	-0.521, 0.390	0.778	0.766, 0.817	0.792	4675	0.276	-0.155, 0.707	0.209	0.183, 0.234	0.208	5290
16. Anxiety	Beta	-0.022	-0.023, 0.019	0.849	0.805, 0.853	0.830	4993	-0.022	-0.042, -0.002	0.029	0.017, 0.038	0.026	5355
17. Sleep maintenance	OR	1.001	0.936, 1.071	0.970	0.947, 0.972	0.961	5127	0.983	0.922, 1.048	0.573	0.517, 0.579	0.548	5488

Note. The intergenerational analysis represents offspring phenotypes <10 years regressed on maternal PRS. The childhood analysis represents offspring phenotypes <10 years regressed on offspring PRS.

Appendix G

G1 Associations between the maternal and offspring smoking initiation PRS and phenotypes in mothers during and outside of pregnancy and adolescence.

		Regression analyses			Permutation testing		
Phenotype	**Effect estimate**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	Sample size
			Mothers outside of pregnancy				
Mental health							
Depression symptoms	OR	1.070	0.969, 1.182	0.161	0.117, 0.161	0.138	4725
Anxiety symptoms	OR	1.028	0.934, 1.131	0.542	0.515, 0.577	0.546	4740
Bulimia	OR	1.081	0.926, 1.261	0.295	0.319, 0.379	0.349	6799
Drug addiction	OR	0.938	0.594, 1.480	0.764	0.748, 0.801	0.775	6799
Alcoholism	OR	1.243	0.903, 1.711	0.163	0.118, 0.162	0.139	6799
Schizophrenia	OR	0.839	0.386, 1.825	0.632	0.585, 0.646	0.616	6799
Anorexia nervosa	OR	1.062	0.886, 1.272	0.484	0.429, 0.491	0.460	6799
Severe depression	OR	1.178	1.064, 1.303	0.004	<0.001, 0.004	<0.001	6799
Other psychological problem	OR	1.146	0.941, 1.396	0.157	0.103, 0.145	0.123	6799
Substance use							
Alcohol							
Alcohol drinking before pregnancy	OR	1.129	1.017, 1.253	0.026	0.006, 0.020	0.011	7199
Binge drinking	Beta	0.050	0.020, 0.080	0.001	<0.001, 0.004	<0.001	4866
Daily alcohol units at child age 4	Beta	0.023	0.003, 0.044	0.027	0.018, 0.039	0.027	5680
Daily alcohol units at child age 8	Beta	-0.003	-0.033, 0.027	0.838	0.799, 0.847	0.824	2707
AUDIT score	Beta	0.023	0.002, 0.045	0.036	0.034, 0.061	0.046	2424
Caffeine							
Total caffeine consumption	Beta	8.568	4.948, 12.187	<0.001	<0.001, 0.004	<0.001	4783
Non-mental health							
Life events	Beta	0.021	-0.012, 0.055	0.212	0.222, 0.277	0.249	4219
Sleep duration	Beta	-0.023	-0.049, 0.004	0.099	0.088, 0.127	0.106	1867
Impulsivity personality trait	Beta	0.072	-0.034, 0.177	0.183	0.147, 0.195	0.170	4847
Monotony avoidance personality trait	Beta	0.242	0.099, 0.386	0.001	<0.001, 0.006	0.001	4794
Anger personality trait	Beta	0.341	0.207, 0.475	<0.001	<0.001, 0.004	<0.001	4769
Suspicion personality trait	Beta	0.125	0.016, 0.234	0.024	0.012, 0.031	0.020	4856
Detachment personality trait	Beta	-0.058	-0.169, 0.053	0.304	0.293, 0.352	0.322	4753
Physical activity	OR	0.933	0.858, 1.014	0.094	0.040, 0.069	0.053	2787

Social class	Beta	0.020	-0.024, 0.064	0.379	0.422, 0.484	0.453	2906
Education	Beta	-0.092	-0.124, -0.060	2.19×10^{-8}	<0.001, 0.004	<0.001	4919
BMI before pregnancy	Beta	0.210	0.117, 0.302	9.67×10^{-6}	<0.001, 0.004	<0.001	6398
Image perception before pregnancy	Beta	0.055	0.029, 0.082	3.26×10^{-5}	<0.001, 0.004	<0.001	6623
Total number of outcomes tested = 27							
Mothers during pregnancy							
Mental health							
Depression (18 wks)	OR	1.115	1.028, 1.211	0.013	0.001, 0.010	0.004	6734
Depression (32 wks)	OR	1.124	1.039, 1.216	0.007	<0.001, 0.006	0.001	6751
Anxiety	OR	0.999	0.915, 1.091	0.991	0.979, 0.994	0.988	6645
Hypersensitivity to interpersonal rejection	Beta	-0.474	-0.846, -0.102	0.012	0.003, 0.016	0.008	7167
Feelings becoming a parent	Beta	-0.003	-0.025, 0.018	0.752	0.724, 0.778	0.752	7165
Substance use							
Caffeine							
Total caffeine (18wks)	Beta	7.352	4.748, 9.957	3.25×10^{-8}	<0.001, 0.004	<0.001	7220
Total caffeine (32wks)	Beta	6.282	3.693, 8.872	2.02×10^{-6}	<0.001, 0.004	<0.001	6767
Alcohol							
Binge drinking (18wks)	Beta	0.043	0.024, 0.061	8.07×10^{-6}	<0.001, 0.004	<0.001	7171
Binge drinking (32wks)	Beta	0.034	0.014, 0.054	0.001	<0.001, 0.004	<0.001	5324
Weekly alcohol units (32wks)	Beta	0.160	0.033, 0.286	0.013	0.003, 0.014	0.007	4294
Other substances							
Cannabis use in pregnancy	OR	1.165	0.977, 1.389	0.082	0.046, 0.077	0.060	6918
Hard drug use in pregnancy	OR	0.990	0.568, 1.726	0.971	0.947, 0.972	0.961	7147
Non-mental health							
Education	Beta	-0.100	-0.128, -0.071	1.01×10^{-11}	<0.001, 0.004	<0.001	6954
Social class	Beta	0.050	0.023, 0.078	3.19×10^{-4}	<0.001, 0.004	<0.001	5854
Life events in pregnancy	Beta	0.046	0.018, 0.074	0.001	<0.001, 0.004	<0.001	6744
Image perception in pregnancy	Beta	0.145	0.045, 0.245	0.005	0.004, 0.017	0.009	6699
Image perception change	Beta	0.077	-0.011, 0.166	0.087	0.079, 0.117	0.097	6549
Activity level compared with other pregnant women	Beta	0.011	-0.008, 0.029	0.262	0.226, 0.281	0.253	6611
Physical activity	OR	1.007	0.951, 1.066	0.795	0.764, 0.816	0.795	6767
Vomiting in first three months of pregnancy	OR	0.979	0.927, 1.034	0.418	0.367, 0.428	0.397	6797
Sleep problems 18 weeks gestation	Beta	0.019	0.001, 0.036	0.036	0.015, 0.036	0.005	6742
Sleep problems 32 weeks gestation	Beta	0.028	0.009, 0.046	0.003	<0.001, 0.004	<0.001	6743
Total number of outcomes tested = 22							

Offspring: Adolescents							
Mental health							
Conduct disorder symptoms	Beta	0.041	0.015, 0.066	0.002	<0.001, 0.006	0.001	3834
ADHD symptoms	Beta	0.050	0.016, 0.085	0.004	0.001, 0.010	0.004	3852
Oppositional-defiant disorder symptoms	Beta	0.036	0.011, 0.060	0.004	<0.001, 0.006	0.001	3436
Psychosis positive symptoms age 12	Beta	0.015	<0.001, 0.029	0.046	0.034, 0.061	0.046	4974
Psychosis negative symptoms age 16	Beta	-0.022	-0.059, 0.015	0.251	0.230, 0.285	0.257	3511
Psychosis positive symptoms age 18	Beta	0.012	-0.004, 0.028	0.134	0.089, 0.129	0.134	3403
PTSD disorder	Beta	0.013	-0.002, 0.028	0.085	0.042, 0.071	0.055	4008
Depression score age 17 (MFQ)	Beta	0.008	-0.004, 0.020	0.178	0.164, 0.214	0.188	3212
Depression symptom score age 18 (CIS-R)	Beta	0.015	-0.010, 0.041	0.236	0.210, 0.264	0.236	3303
Eating disorder age 16	Beta	-0.002	-0.007, 0.002	0.281	0.247, 0.303	0.274	3543
Eating disorder age 13	Beta	-0.001	-0.003, 0.001	0.395	0.452, 0.514	0.483	4256
Specific phobia symptoms	Beta	0.010	-0.008, 0.028	0.298	0.269,0.326	0.297	3293
Emotional problems symptoms	Beta	-0.009	-0.032, 0.014	0.422	0.410, 0.472	0.441	4073
Self-harming behaviour	OR	0.958	0.810, 1.135	0.596	0.562, 0.624	0.593	2576
Depression symptoms score age 14 (MFQ)	Beta	-0.002	-0.016, 0.013	0.821	0.808, 0.856	0.833	4574
Anxiety score	Beta	0.002	-0.023, 0.027	0.848	0.830, 0.874	0.853	3293
Total behavioural difficulties score	Beta	0.036	-0.001, 0.073	0.055	0.035, 0.062	0.047	4055
Substance use							
Cannabis use	OR	1.225	1.127, 1.330	<0.001	<0.001, 0.004	<0.001	3571
AUDIT risk score age 18	Beta	0.041	0.018, 0.064	0.001	<0.001, 0.004	<0.001	3008
Binge drinking age 18	Beta	0.070	0.025, 0.114	0.002	0.001, 0.010	0.004	2829
AUDIT total score age 18	Beta	0.069	0.029, 0.109	6.58×10^{-4}	<0.001, 0.004	<0.001	3008
Number of alcoholic drinks on a typical day	Beta	0.066	0.022, 0.111	0.003	<0.001, 0.007	0.002	2826
Number of drinks to feel tipsy	Beta	0.050	0.008, 0.093	0.020	0.017, 0.038	0.026	2391
Number of drinks to feel different after first five times drinking	Beta	0.097	-0.016, 0.211	0.093	0.073, 0.109	0.090	299
Binge drinking age 13	Beta	0.039	-0.065, 0.142	0.461	0.421, 0.483	0.452	464
Frequency of having alcoholic drinks	Beta	0.008	-0.024, 0.039	0.641	0.586, 0.647	0.617	3626
Number of times had whole drink age 13	Beta	0.007	-0.066, 0.079	0.860	0.838, 0.882	0.861	1103
Frequency of cannabis smoking	Beta	-0.016	-0.092, 0.060	0.676	0.625, 0.684	0.655	1035
Total caffeine age 13	Beta	0.008	-0.030, 0.046	0.680	0.639, 0.698	0.669	3405
Non-mental health							
BMI	Beta	0.239	0.104, 0.373	0.001	<0.001, 0.004	<0.001	3606
IQ	Beta	-0.582	-1.006, -0.159	0.007	0.003, 0.016	0.008	3720

GCSE grades D-G	OR	1.086	0.988, 1.193	0.083	0.040, 0.069	0.053	2182
GCSE grades A-C	OR	0.819	0.651, 1.032	0.085	0.063, 0.097	0.079	2360
Extraversion personality trait	Beta	0.362	0.155, 0.569	0.001	<0.001, 0.006	0.001	4354
Conscientiousness personality trait	Beta	-0.217	-0.403, -0.031	0.022	0.006, 0.021	0.012	4162
Emotional Stability personality trait	Beta	-0.073	-0.263, 0.117	0.449	0.436, 0.498	0.467	4224
Intellect personality trait	Beta	-0.053	-0.226, 0.119	0.545	0.491, 0.553	0.522	4263
Agreeableness personality trait	Beta	-0.037	-0.188, 0.113	0.628	0.610, 0.671	0.641	4279
Sleep maintenance	Beta	0.025	<0.001, 0.051	0.051	0.039, 0.068	0.052	3418
Sleep initiation (time to fall asleep)	Beta	0.008	-0.027, 0.043	0.641	0.616, 0.677	0.647	3626
Sleep duration (hours of sleep)	Beta	0.015	-0.017, 0.047	0.360	0.304, 0.363	0.333	3726
Frequency of doing exercise	Beta	-0.004	-0.027, 0.020	0.762	0.736, 0.790	0.764	4270
Life events	Beta	-0.006	-0.043, 0.031	0.756	0.762, 0.814	0.789	3376
Total number of outcomes tested = 44							

Appendix H

H1 Associations between maternal and offspring lifetime smoking PRS and phenotypes in mothers during and outside pregnancy and adolescence.

			Regression analyses		Permutation testing		
Phenotype	**Effect estimate**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	**Sample size**
Mothers outside of pregnancy							
Mental health							
Depression symptoms	OR	1.007	0.913, 1.109	0.886	0.876, 0.915	0.897	4725
Anxiety symptoms	OR	1.005	0.915, 1.105	0.904	0.880, 0.918	0.900	4740
Bulimia	OR	1.140	0.964, 1.349	0.114	0.083, 0.121	0.101	6799
Drug addiction	OR	0.983	0.596, 1.620	0.941	0.912, 0.945	0.930	6799
Alcoholism	OR	1.270	0.968, 1.667	0.079	0.101, 0.143	0.121	6799
Schizophrenia	OR	1.585	0.870, 2.889	0.120	0.197, 0.249	0.222	6799
Anorexia Nervosa	OR	1.150	0.932, 1.419	0.174	0.096, 0.136	0.115	6799
Severe depression	OR	1.159	1.049, 1.280	0.007	0.001, 0.010	0.004	6799
Other psychiatric problem	OR	1.156	0.949, 1.408	0.134	0.079, 0.117	0.097	6799
Substance use							
Alcohol							
Alcohol drinking before pregnancy	OR	1.010	0.910, 1.122	0.833	0.816, 0.862	0.840	7199
Binge drinking	Beta	0.039	0.009, 0.068	0.010	<0.001, 0.004	<0.001	4867
Daily alcohol units at child age 4	Beta	0.028	0.007, 0.049	0.008	0.006, 0.020	0.011	5680
Daily alcohol units at child age 8	Beta	-0.009	-0.039, 0.021	0.559	0.564, 0.626	0.595	2707
AUDIT score	Beta	0.014	-0.007, 0.035	0.181	0.174, 0.224	0.198	2424
Caffeine							
Total caffeine consumption	Beta	8.698	5.083, 12.313	2.46×10^{-6}	<0.001, 0.004	<0.001	4783
Non-mental health							
Life events	Beta	0.026	-0.009, 0.060	0.142	0.117, 0.161	0.138	4219
Sleep duration	Beta	-0.018	-0.046, 0.009	0.183	0.163, 0.213	0.187	1867
Impulsivity personality trait	Beta	0.111	0.006, 0.217	0.039	0.033, 0.060	0.045	4847
Monotony avoidance personality trait	Beta	0.184	0.037, 0.332	0.014	0.008, 0.025	0.015	4794
Anger personality trait	Beta	0.246	0.115, 0.377	2.34×10^{-4}	<0.001, 0.004	<0.001	4769
Suspicion personality trait	Beta	0.164	0.057, 0.272	0.003	0.003, 0.016	0.008	4856
Detachment personality trait	Beta	-0.061	-0.175, 0.054	0.301	0.258, 0.315	0.286	4753
Physical activity	OR	0.996	0.915, 1.085	0.929	0.919, 0.950	0.936	2787

Social class	Beta	0.028	-0.015, 0.072	0.204	0.185, 0.237	0.210	2906
Education	Beta	-0.083	-0.115, -0.051	4.33×10^{-7}	<0.001, 0.004	<0.001	4919
BMI before pregnancy	Beta	0.160	0.065, 0.254	0.001	0.001, 0.010	0.004	6398
Image perception before pregnancy	Beta	0.028	0.002, 0.054	0.037	0.029, 0.054	0.040	6623
Total number of outcomes tested = 27							
Mothers during pregnancy							
Mental health							
Depression (18wks)	OR	1.076	0.997, 1.163	0.060	0.034, 0.061	0.046	6734
Depression (32wks)	OR	1.078	0.999, 1.164	0.053	0.015, 0.036	0.024	6751
Anxiety (18 wks)	OR	1.064	0.980, 1.155	0.127	0.087, 0.126	0.105	6645
Hypersensitivity to interpersonal rejection	Beta	-0.296	-0.657, 0.065	0.108	0.097, 0.137	0.116	7167
Feelings becoming a parent	Beta	-0.012	-0.034, 0.009	0.266	0.236, 0.291	0.263	7165
Substance use							
Caffeine							
Total caffeine (18wks)	Beta	6.759	4.239, 9.280	<0.001	<0.001, 0.004	<0.001	7220
Total caffeine (32wks)	Beta	5.325	2.776, 7.874	<0.001	<0.001, 0.004	<0.001	6767
Alcohol							
Binge drinking (18wks)	Beta	0.024	0.005, 0.042	0.012	0.003, 0.016	0.008	7171
Binge drinking (32wks)	Beta	0.020	0.001, 0.039	0.044	0.049, 0.080	0.063	5324
Weekly alcohol units (32wks)	Beta	0.133	0.034, 0.233	0.009	0.008, 0.025	0.015	4294
Other substances							
Cannabis use during pregnancy	OR	1.106	0.942, 1.299	0.197	0.175, 0.225	0.199	6918
Hard drugs	OR	1.053	0.670, 1.653	0.809	0.779, 0.829	0.805	7147
Non-mental health							
Education	Beta	-0.094	-0.122, -0.065	<0.001	<0.001, 0.004	<0.001	6954
Social class	Beta	0.064	0.037, 0.091	<0.001	<0.001, 0.004	<0.001	5854
Life events during pregnancy	Beta	0.018	-0.010, 0.045	0.214	0.196, 0.248	0.221	6744
Image perception during pregnancy	Beta	0.121	0.023, 0.219	0.016	0.011, 0.028	0.018	6699
Image perception change	Beta	0.004	-0.087, 0.095	0.931	0.906, 0.940	0.924	6549
Activity level compared with other pregnant women	Beta	-0.001	-0.019, 0.017	0.911	0.892, 0.928	0.911	6611
Physical activity	OR	1.002	0.946, 1.061	0.952	0.941, 0.968	0.956	6767
Vomited first three months in pregnancy	OR	0.979	0.928, 1.033	0.412	0.373, 0.435	0.404	6797
Sleep (18 wks)	Beta	0.014	-0.003, 0.032	0.108	0.089, 0.129	0.108	6742
Sleep (32 wks)	Beta	0.034	0.016, 0.052	<0.001	<0.001, 0.004	<0.001	6743
Offspring: Adolescence							
Mental health							
Conduct disorder symptoms	Beta	0.057	0.031, 0.082	<0.001	<0.001, 0.004	<0.001	3834
Psychosis positive symptoms age 12	Beta	0.024	0.010, 0.018	0.001	<0.001, 0.004	<0.001	4974

Depression symptoms age 17 (MFQ)	Beta	0.019	0.007, 0.036	0.002	<0.001, 0.006	0.001	3212
Total behavioural difficulties	Beta	0.055	0.019, 0.091	0.003	0.001, 0.009	0.003	4055
Psychosis positive symptoms age 18	Beta	0.020	0.004, 0.018	0.014	0.004, 0.017	0.009	3403
ADHD symptoms	Beta	0.031	-0.003, 0.030	0.075	0.078, 0.116	0.096	3852
Eating disorder age 16	Beta	0.003	-0.001, 0.041	0.146	0.138, 0.184	0.160	3543
Depression symptoms score age 17	Beta	0.016	-0.010, 0.039	0.228	0.218, 0.272	0.244	3303
Specific phobia symptoms	Beta	0.010	-0.009, 0.031	0.301	0.261, 0.318	0.289	3293
PTSD symptoms	Beta	0.006	-0.007, 0.027	0.360	0.340, 0.401	0.370	4008
Oppositional defiant disorder	Beta	0.011	-0.013, 0.169	0.374	0.342, 0.403	0.372	3436
Anxiety symptoms score	Beta	0.012	-0.015, 0.066	0.387	0.345, 0.406	0.375	3293
Eating disorder age 13	Beta	0.001	-0.001, 0.007	0.475	0.454, 0.516	0.485	4256
Depression symptoms age 14 (MFQ)	Beta	0.004	-0.010, 0.019	0.573	0.568, 0.630	0.599	4574
Emotional problems symptoms	Beta	0.004	-0.018, 0.036	0.700	0.655, 0.714	0.685	4073
Psychosis negative symptoms age 16	Beta	0.002	-0.034, 0.038	0.922	0.892, 0.928	0.911	3511
Self-harming behaviour	OR	0.994	0.834, 1.185	0.944	0.927, 0.957	0.943	2576
Substance use							
Alcohol							
Number of drinks needed to feel different	Beta	0.099	-0.020, 0.218	0.103	0.101, 0.143	0.121	299
Binge drinking age 13	Beta	0.090	-0.017, 0.197	0.099	0.073, 0.109	0.090	464
Number of times had a whole drink past 6 months	Beta	-0.007	-0.076, 0.062	0.840	0.825, 0.871	0.849	1103
Number of alcoholic drinks on a typical day	Beta	0.039	-0.005, 0.082	0.083	0.058, 0.091	0.073	2826
Binge drinking age 18	Beta	0.063	0.017, 0.109	0.007	0.002, 0.012	0.005	2829
Frequency of having alcoholic drinks	Beta	-0.002	-0.032, 0.027	0.876	0.833, 0.877	0.856	2886
AUDIT risk score age 18	Beta	0.038	0.014, 0.061	0.002	<0.001, 0.006	0.001	3008
AUDIT total score age 18	Beta	0.035	-0.004, 0.075	0.082	0.088, 0.127	0.106	3008
Number of drinks needed to feel tipsy	Beta	0.047	0.004, 0.089	0.032	0.030, 0.055	0.041	2391
Tobacco							
Cannabis use	OR	1.074	0.990, 1.164	0.082	0.040, 0.069	0.053	3571
Frequency of cannabis use	Beta	0.042	-0.031, 0.115	0.261	0.248, 0.305	0.276	1035
Caffeine							
Total caffeine consumption	Beta	0.018	-0.019, 0.055	0.348	0.342, 0.403	0.372	3405
Non-mental health							
Extraversion personality trait	Beta	0.445	0.244, 0.646	<0.001	<0.001, 0.004	<0.001	4354
Conscientiousness personality trait	Beta	-0.187	-0.367, -0.008	0.041	0.030, 0.056	0.042	4162
Agreeableness personality trait	Beta	0.010	-0.132, 0.151	0.893	0.896, 0.932	0.915	4279
Intellect personality trait	Beta	-0.009	-0.174, 0.156	0.918	0.900, 0.935	0.919	4263
Emotional stability personality trait	Beta	-0.065	-0.253, 0.124	0.501	0.489, 0.551	0.520	4224
IQ	Beta	-0.741	-1.163, -0.320	0.001	<0.001, 0.004	<0.001	3720

BMI	Beta	0.205	0.078, 0.331	0.001	0.001, 0.010	0.004	3606
Sleep maintenance	Beta	0.033	0.008, 0.059	0.011	0.005, 0.018	0.010	3418
GCSE grades D-G	OR	1.106	1.008, 1.213	0.036	0.011, 0.028	0.018	2182
Frequency of doing exercise	Beta	-0.025	-0.048, -0.001	0.039	0.032, 0.059	0.044	4270
Sleep duration (hours of sleep)	Beta	-0.025	-0.055, 0.005	0.097	0.088, 0.127	0.106	3726
Sleep initiation (time to fall asleep)	Beta	0.025	-0.009, 0.060	0.150	0.127, 0.173	0.149	3626
GCSE grades A-C	OR	0.840	0.652, 1.082	0.160	0.101, 0.142	0.120	2360
Life events	Beta	0.011	-0.026, 0.047	0.570	0.583, 0.644	0.614	3376
Total number of outcomes tested = 22							

Appendix I

I1 Associations between caffeine PRS and phenotypes in mothers during and outside of pregnancy and adolescence

			Regression analyses		Permutation testing		
Phenotype	**Effect estimate**	**Effect size**	**95% CI**	**P-value**	**95% CI**	**P-value**	**Sample size**
Mothers outside of pregnancy							
Mental health							
Depression symptoms	OR	1.040	0.943, 1.148	0.398	0.507, 0.569	0.538	4725
Anxiety symptoms	OR	0.979	0.890, 1.077	0.641	0.340, 0.401	0.370	4740
Bulimia	OR	1.105	0.932, 1.309	0.225	0.180, 0.231	0.205	6799
Drug addiction	OR	0.987	0.657, 1.481	0.943	0.941, 0.968	0.956	6799
Alcoholism	OR	0.947	0.697, 1.287	0.706	0.697, 0.753	0.726	6799
Schizophrenia	OR	0.434	0.244, 0.772	0.008	0.021, 0.044	0.031	6799
Anorexia Nervosa	OR	1.076	0.899, 1.289	0.390	0.354, 0.415	0.384	6799
Severe depression	OR	1.054	0.953, 1.166	0.277	0.218, 0.272	0.244	6799
Other psychiatric problem	OR	1.048	0.867, 1.266	0.601	0.529, 0.591	0.560	6799
Substance use							
Tobacco							
Ever smoking	OR	1.010	0.959, 1.064	0.679	0.626. 0.685	0.656	7194
Number of cigarettes smoked past 2 weeks	Beta	0.304	-0.271, 0.879	0.300	0.270, 0.327	0.298	845
Number of cigarettes smoked before pregnancy	Beta	0.042	-0.028, 0.111	0.245	0.224, 0.279	0.251	3426
Alcohol							
Alcohol drinking before pregnancy	OR	0.972	0.876, 1.078	0.558	0.016, 0.514	0.545	7199
Binge drinking	Beta	0.004	-0.024, 0.032	0.786	0.752, 0.804	0.779	4867
Daily alcohol units at child age 4	Beta	0.007	-0.013, 0.027	0.518	0.467, 0.529	0.498	5680
Daily alcohol units at child age 8	Beta	0.014	-0.016, 0.044	0.347	0.354, 0.415	0.384	2707
AUDIT score	Beta	0.007	-0.013, 0.028	0.473	0.450, 0.512	0.481	2424
Non-mental health							
Life events	Beta	-0.025	-0.059, 0.008	0.141	0.115, 0.159	0.136	4219
Sleep duration	Beta	-0.007	-0.034, 0.019	0.588	0.535, 0.597	0.566	1867
Impulsivity personality trait	Beta	0.042	-0.063, 0.146	0.436	0.397, 0.459	0.428	4847
Monotony avoidance personality trait	Beta	-0.107	-0.251, 0.037	0.144	0.131, 0.177	0.153	4794
Anger personality trait	Beta	0.014	-0.115, 0.144	0.830	0.784, 0.834	0.810	4769
Suspicion personality trait	Beta	-0.017	-0.128, 0.095	0.772	0.718, 0.773	0.746	4856
Detachment personality trait	Beta	0.012	-0.102, 0.125	0.841	0.815, 0.861	0.839	4753
Physical activity	OR	0.966	0.890, 1.050	0.387	0.015, 0.340	0.370	2787
Social class	Beta	-0.009	-0.053, 0.035	0.696	0.680, 0.737	0.709	2906

Education	Beta	0.002	-0.030, 0.035	0.889	0.882, 0.920	0.902	4919
BMI before pregnancy	Beta	0.083	-0.008, 0.174	0.075	0.058, 0.091	0.073	6398
Image perception before pregnancy	Beta	-0.003	-0.029, 0.023	0.820	0.796, 0.844	0.821	6623
Total number of outcomes tested = 29							
Mothers during pregnancy							
Mental health							
Depression symptoms (18 wks)	OR	0.982	0.905, 1.067	0.647	0.616, 0.677	0.647	6734
Depression symptoms (32 wks)	OR	0.994	0.920, 1.074	0.870	0.823, 0.869	0.847	6751
Anxiety symptoms	OR	1.005	0.919, 1.099	0.910	0.891, 0.927	0.910	6645
Hypersensitivity to interpersonal rejection	Beta	-0.040	0.838, 0.882	0.833	0.838, 0.882	0.861	7167
Feelings becoming a parent	Beta	0.007	0.496, 0.558	0.528	0.496, 0.558	0.527	7165
Substance use							
Tobacco							
Ever smoked in pregnancy	OR	1.007	0.947, 1.070	0.813	0.783, 0.833	0.809	6718
Smoking first three months in pregnancy	OR	1.029	0.966, 1.097	0.343	0.275, 0.333	0.303	7237
Caffeine							
Reduced caffeine consumption during pregnancy	OR	1.054	1.001, 1.111	0.046	0.017, 0.038	0.026	7269
Reduced coffee consumption during pregnancy	OR	1.061	1.008, 1.117	0.028	0.006, 0.021	0.012	7269
Stopped drinking cola during pregnancy	OR	1.094	0.991, 1.208	0.072	0.034, 0.061	0.046	4570
Never drank coffee	OR	1.064	0.988, 1.146	0.093	0.059, 0.092	0.074	6782
Never drank cola	OR	1.002	0.946, 1.062	0.933	0.915, 0.947	0.932	6744
Stopped drinking coffee during pregnancy	OR	1.039	0.981, 1.100	0.175	0.104, 0.146	0.124	5809
Never has been drinking caffeine	OR	0.967	0.918, 1.018	0.179	0.126, 0.170	0.147	7269
Stopped drinking tea during pregnancy	OR	1.039	0.973, 1.109	0.226	0.184, 0.236	0.209	6082
Reduced cola consumption during pregnancy	OR	1.039	0.967, 1.116	0.272	0.244, 0.300	0.271	7269
Never drank tea	OR	1.041	0.950, 1.141	0.353	0.269, 0.326	0.297	6754
Reduced tea consumption during pregnancy	OR	1.026	0.970, 1.085	0.333	0.290, 0.349	0.319	7269
Consumed more caffeine during pregnancy	OR	1.038	0.948, 1.137	0.384	0.333, 0.394	0.363	7269
No change in caffeine consumption during pregnancy	OR	0.986	0.934, 1.041	0.580	0.532, 0.594	0.563	7269
Craved or had more caffeine during pregnancy	OR	0.985	0.909, 1.068	0.698	0.682, 0.739	0.711	7269
Craved or had more coffee during pregnancy	OR	0.977	0.824, 1.160	0.774	0.700, 0.756	0.729	6782
Craved or had more tea during pregnancy	OR	1.013	0.927, 1.108	0.750	0.718, 0.773	0.746	6754
Alcohol							
	Beta		-0.041, -0.003				
Binge drinking (32wks)		-0.022		0.026	0.031, 0.057	0.043	5324
Binge drinking (18wks)	Beta	-0.010	-0.029, 0.008	0.268	0.249, 0.306	0.277	7171
Weekly alcohol units (32wks)	Beta	-0.056	-0.167, 0.056	0.329	0.314, 0.373	0.343	4294
Craved or had more alcohol during pregnancy	OR	0.949	0.569, 1.584	0.828	0.826, 0.872	0.850	6771

Other substances							
Cannabis use in first three months during pregnancy	OR	1.124	0.952, 1.328	0.151	0.134, 0.180	0.156	6918
Hard drugs during pregnancy	OR	0.995	0.664, 1.491	0.980	0.968, 0.987	0.979	7147
Non-mental health							
Life events during pregnancy	Beta	0.001	0.154, 0.202	0.944	0.154, 0.202	0.177	6930
Activity level compared with other pregnant women	Beta	0.011	0.227, 0.282	0.234	0.227, 0.282	0.254	6611
Image perception during pregnancy	Beta	-0.031	0.517, 0.579	0.540	0.517, 0.579	0.548	6699
Physical activity	Beta	-0.003	0.754, 0.806	0.780	0.754, 0.806	0.781	6767
Social class	Beta	0.028	0.035, 0.062	0.043	0.035, 0.062	0.047	6954
Image perception change	Beta	0.019	0.394, 0.456	0.678	0.394, 0.456	0.425	3741
Education	Beta	-0.005	0.658, 0.717	0.709	0.658, 0.717	0.688	6954
Vomiting in first three months during pregnancy	OR	1.003	0.950, 1.059	0.903	0.871, 0.911	0.892	0.903
Sleeping problems (18 wks)	Beta	0.002	0.797, 0.845	0.825	0.797, 0.845	0.822	6742
Sleeping problems (32 wks)	Beta	-0.003	0.726, 0.780	0.733	0.726, 0.780	0.754	6743
Total number of outcomes tested = 39							
Offspring: Adolescence							
Mental health							
Conduct disorder symptoms	Beta	0.012	-0.014, 0.039	0.362	0.332, 0.393	0.362	3834
Depression symptoms score age 18	Beta	0.013	-0.013, 0.039	0.314	0.288, 0.347	0.317	3303
Specific phobia symptoms	Beta	0.001	-0.019, 0.020	0.937	0.935, 0.963	0.950	3293
Emotional problems score	Beta	-0.022	-0.047, 0.002	0.072	0.063, 0.097	0.079	3593
Anxiety symptoms	Beta	0.001	-0.024, 0.026	0.913	0.899, 0.934	0.918	3293
Eating disorder age 13	Beta	-0.001	-0.003, 0.001	0.289	0.346, 0.407	0.376	4256
Eating disorder age 16	Beta	0.003	-0.001, 0.007	0.184	0.166, 0.216	0.190	3543
ADHD symptoms	Beta	-0.027	-0.065, 0.010	0.146	0.142, 0.188	0.164	3435
Depression score age 14 (MFQ)	Beta	-0.002	-0.016, 0.013	0.835	0.817, 0.863	0.841	4574
Depression score age 17 (MFQ)	Beta	-0.003	-0.015, 0.010	0.685	0.647, 0.706	0.677	3212
Psychosis negative symptoms age 16	Beta	0.000	-0.037, 0.036	0.996	0.993, 1.000	0.998	3511
Total behavioural difficulties	Beta	-0.017	-0.056, 0.022	0.397	0.364, 0.425	0.394	3603
Psychosis positive symptoms age 12	Beta	0.009	-0.006, 0.024	0.230	0.198, 0.250	0.223	4974
Psychosis positive symptoms age 18	Beta	0.011	-0.006, 0.027	0.200	0.097, 0.137	0.116	3403
PTSD disorder symptoms	Beta	-0.013	-0.027, 0.002	0.091	0.052, 0.084	0.067	4008
Self-harming behaviour	OR	0.985	0.811, 1.196	0.869	0.836, 0.880	0.859	2576
	Beta	-0.011	-0.036, 0.013	0.367	0.561, 0.623	0.592	3436
Oppositional-defiant disorder symptoms age 15							
Substance use							
Tobacco							

Age when first smoked a cigarette	Beta	-0.013	-0.056, 0.030	0.553	0.535, 0.597	0.566	1064
Has smoked a cigarette	OR	1.045	0.927, 1.179	0.443	0.384, 0.446	0.415	2089
Total number of cigarettes smoked age 14	OR	0.990	0.767, 1.277	0.931	0.900, 0.935	0.919	461
Total number of cigarettes smoked age 18	Beta	0.069	-0.023, 0.162	0.142	0.114, 0.158	0.135	1144
Alcohol							
Number of drinks to feel different	Beta	-0.039	-0.154, 0.076	0.505	0.494, 0.556	0.525	299
Binge drinking age 13	Beta	0.010	-0.094, 0.113	0.854	0.834, 0.878	0.857	464
Number of times had whole drink age 13	Beta	0.012	-0.059, 0.083	0.748	0.700, 0.756	0.729	1103
Number of alcoholic drinks on a typical day	Beta	-0.012	-0.058, 0.034	0.609	0.580, 0.641	0.611	2826
Binge drinking age 18	Beta	0.010	-0.036, 0.056	0.670	0.632, 0.691	0.662	2829
Frequency having alcoholic drinks	Beta	0.011	-0.020, 0.042	0.485	0.462, 0.524	0.493	2886
AUDIT risk score age 18	Beta	-0.007	-0.030, 0.017	0.562	0.539, 0.601	0.570	3008
AUDIT total score age 18	Beta	0.010	-0.031, 0.050	0.647	0.986, 0.997	0.993	3008
Number of drinks needed to feel tipsy	Beta	-0.015	-0.059, 0.029	0.500	0.461, 0.523	0.492	2391
Other substances							
Cannabis use	OR	0.977	0.900, 1.060	0.551	0.494, 0.556	0.525	3571
Frequency of cannabis use	Beta	0.018	-0.057, 0.093	0.636	0.613, 0.674	0.644	1035
Non-mental health							
BMI	Beta	0.031	-0.100, 0.161	0.645	0.612, 0.673	0.643	3606
Agreeableness personality trait	Beta	0.066	-0.080, 0.211	0.376	0.368, 0.430	0.399	4279
Conscientiousness personality trait	Beta	-0.044	-0.218, 0.130	0.617	0.600, 0.661	0.631	4162
Intellect personality trait	Beta	0.100	-0.069, 0.269	0.245	0.223, 0.278	0.250	4263
Emotional stability personality trait	Beta	-0.067	-0.263, 0.130	0.506	0.472, 0.534	0.503	4224
Extraversion personality trait	Beta	-0.042	-0.243, 0.159	0.682	0.657, 0.716	0.687	4354
Frequency of doing exercise	Beta	-0.009	-0.032, 0.014	0.450	0.443, 0.505	0.474	4270
Sleep duration (hours of sleep)	Beta	-0.016	-0.047, 0.014	0.294	0.260, 0.317	0.288	3726
GCSE grades A-C	OR	1.467	1.146, 1.877	0.005	<0.001, 0.004	<0.001	2360
GCSE grades D-G	OR	1.007	0.914, 1.109	0.876	0.846, 0.889	0.869	2182
IQ	Beta	0.138	-0.293, 0.569	0.531	0.497, 0.559	0.528	3720
Sleep initiation	Beta	0.015	-0.019, 0.050	0.385	0.353, 0.414	0.383	3626
Sleep maintenance	Beta	-0.003	-0.028, 0.022	0.804	0.813, 0.859	0.837	3418
Life events	Beta	-0.007	-0.045, 0.031	0.733	0.690, 0.747	0.719	3376
Total number of outcomes tested = 46							

Appendix J

J1 Correlation between smoking, caffeine and alcohol PRS

	Smoking initiation PRS	**Lifetime smoking PRS**	**Caffeine PRS**	**Alcohol PRS**
Smoking initiation PRS	-	0.35	0.01	0.08
-	**Lifetime smoking PRS**	-	-0.01	0.02
-	-	**Caffeine PRS**	-	0.12

Appendix K

K1 ALSPAC definition of variables and covariates

- **Maternal coffee consumption:** Coffee intake was initially assessed in cups per weekday and cups per weekend during 18-weeks gestation. Total cups of coffee were derived by adding up the cups of coffee consumed on weekdays and weekends. Cups of coffee were then transformed to mg/day by: (cups/week*57)/7.
- **Maternal tea consumption:** Tea intake was initially assessed in cups per weekday and cups per weekend during 18-weeks gestation. Total cups of tea were derived by adding up the cups of tea consumed on weekdays and weekends. Cups of tea were then transformed to mg/day by: (cups/week*27)/7.
- **Maternal cola consumption:** Cola intake was initially assessed in cans per weekday and weekend during 18-weeks gestation. Total cans of cola were derived by adding up the cans of cola consumed on weekdays and weekends. Cans were then transformed to mg/day by: (cans/week*20)/7.
- **Maternal total caffeine consumption:** Total maternal caffeine intake during pregnancy in mg/day, summing caffeine from tea, coffee and cola drinks. NAs were treated as 0, unless tea, coffee *and* cola were all missing, in which case the variable was coded as missing.
- **Maternal education** (as proxy for maternal socioeconomic position): Maternal education was assessed in week 32 of gestation and coded as an ordinal variable: "Vocational" = 1, “O level” (at 16, equivalent to lower grades of ordinary-level) = 2, "A level" (ordinary-level school-leaving certificate (at 16) = 3, and “Degree” (advanced-level school-leaving certificate (post-16)/degree) = 4.
- **Maternal smoking during pregnancy** was assessed as an ordinary variable representing 0 = no or early smoking during pregnancy, 1 = Stopped before the second trimester of pregnancy and 2 = Smoking in the third trimester or throughout pregnancy.
- **Maternal age** continuous numeric variable in years assessed at birth of study child.
- **Maternal BMI** was assessed continuous numeric variable in years at 12-weeks gestation**.**
- **Parity** has been assessed at 18-weeks gestation as number of previous pregnancies resulting in either a livebirth or a stillbirth.
- **Gestational age** was calculated (in days) based on the date of the mother’s last menstrual period (LMP) when the mother was certain of this, but for uncertain LMPs and conflicts with clinical assessment the ultrasound assessment was used. Where maternal report and ultrasound assessment conflicted, an experienced obstetrician reviewed clinical records and made a best estimate.
- **Offspring sex** was taken from obstetric records.

K2 Born in Bradford definition of variables and covariates

- **Maternal coffee consumption:** Coffee intake was initially assessed in cups of caffeinated filter/cafetiere coffee and cups of instant coffee consumed per day during 26-28 weeks of gestation. Total cups of coffee were derived by adding up the number of cups of filter/cafetiere coffee and cups of instant coffee consumed per day. Missing data were treated as 0, unless caffeinated filter/cafetiere coffee *and* instant coffee were all missing, in which case the variable was coded as missing. Cups of coffee were then transformed to mg/day by: (cups/week*57)/7.
- **Maternal tea consumption:** Tea intake was initially assessed in cups of caffeinated tea per day during 26-28 weeks of gestation. Cups of tea were then transformed to mg/day by: (cups/week*27)/7.
- **Maternal cola consumption:** Cola intake was assessed in cups of regular, caffeinated cola per day and cups of caffeinated diet cola per day. Missing data were treated as 0, unless caffeinated regular cola *and* diet cola were all missing, in which case the variable was coded as missing Cups were then transformed to mg/day by: (cups/week*20)/7.
- **Maternal total caffeine consumption:** Total maternal caffeine intake during pregnancy in mg/day, summing caffeine from tea, coffee and cola drinks. NAs were treated as 0, unless tea, coffee *and* cola were all missing, in which case the variable was coded as missing.
- **Maternal education** (as proxy for maternal socioeconomic position): Maternal education was assessed in week 26-28 weeks of gestation and coded as an ordinal variable: " <5 GCSE equivalent " = 1, "5 GCSE equivalent" (at 16, equivalent to lower grades of ordinary-level) = 2, " A-level equivalent " (ordinary-level school-leaving certificate (at 16) = 3, and "Higher than A-level" (advanced-level school-leaving certificate (post-16)/degree) = 4.
- **Maternal smoking during pregnancy** was assessed as an ordinary variable representing 0 = no or early smoking during pregnancy, 1 = Stopped before the second trimester of pregnancy and 2 = Smoking in the third trimester or throughout pregnancy.
- **Maternal age** continuous numeric variable in years assessed at 26-28 weeks of gestation.
- **Maternal BMI** was assessed continuous numeric variable in years at 26-28 weeks of gestation**.**
- **Parity** has been assessed at routine healthcare as an integer value.
- **Gestational age** age at completion of questionnaire at weeks 26-28 of gestation (weeks and days).
- **Offspring sex** was assessed at routine healthcare.

K3 MoBa1 definition of variables and covariates

Dr. Christian Page who performed the analysis in MoBa provided the description of variables used in MoBa

- **Maternal coffee consumption:** In MoBa coffee intake was assessed based on type of coffee (instant/espresso coffee) and brewing method (boiled/percolated/filtered) at the 17th week of gestation. Cups were transformed to mg of caffeine based on assuming 85 mg of caffeine per cup of boiled/ percolated/ filtered coffee and 60 mg per cup of instant/espresso coffee
- **Maternal tea consumption**: Maternal cups of tea at 17 weeks gestation were transformed to mg based on the assumption that one cup of tea contains 50mg of caffeine
- **Maternal cola consumption:** Cola consumption was measured as consumption of regular and diet Coca Cola/Pepsi in mugs and the transformed to cups (one mug = two cups; one small bottle = four cups; one large bottle, 1.5L = 12 cups). Cups were then transformed based on the assumption that one cup of cola contains 30mg of caffeine
- **Maternal education** was grouped into 4 levels, 0 = not completed high school, 1 = High school, 2= some college and 4 = four or more years of college/university.
- **Maternal smoking during pregnancy** categorical as 0/1, with 1 for any smoking during pregnancy, and 0 for no-smoking during pregnancy.
- **Maternal pre-pregnancy BMI** self-reported continuous numeric variable in kg/m2 at around 16 weeks gestation**.**
- **Gestational age** was calculated (in days) based on the ultrasound measurements taken at first check up in pregnancy (around 16 weeks, at enrolment in the cohort). If ultrasound measurement was not available, this was based on the last menstrual period.
- **Maternal age** continuous numeric variable in years assessed by the Norwegian Medical Birth Registry
- **Parity** was assessed by the Norwegian Medical Birth Registry as the number of previous live births
- **Offspring sex** was provided by the Norwegian Medical Birth Registry.

K4 Generation R definition of variables and covariates

Giulietta Monasso who performed the analysis in Generation R provided the description of variables used in Generation R

- **Maternal coffee consumption:** Information about maternal coffee intake was assessed by questionnaire at 18-25 weeks gestational. Pregnant women, who indicated to consume coffee, were asked whether they consumed caffeinated or decaffeinated coffee, both, or other. The next question asked about their average number of cups of per day. As the quantity of cups did not differentiate between caffeinated or decaffeinated coffee intake, mothers who indicated to drink both caffeinated and decaffeinated coffee, needed to be excluded (and so were mothers who indicated to drink other types of coffee than caffeinated or decaffeinated). For mothers who reported to drink decaffeinated coffee, the number of cups of coffee per day was set to zero. To calculate the total caffeine intake from coffee, cups were transformed to mg assuming 57 mg of caffeine per cup.
- **Maternal tea consumption:** Women who indicated to consume tea at 18-25 weeks gestational were asked about whether they consumed caffeinated or decaffeinated tea, both, or other. The next question asked about their average number of cups of per day. As the quantity of cups did not differentiate between caffeinated and decaffeinated tea intake, mothers who indicated to drink both caffeinated and decaffeinated tea, needed to be excluded (and so were mothers who indicated to drink other tea than caffeinated or decaffeinated). For mothers who reported to drink decaffeinated tea (herbal or green tea), the number of cups of tea per day was set to zero. To calculate the total caffeine intake from tea, cups were transformed to mg assuming 27 mg of caffeine per cup.
- **Maternal cola consumption:** Not available.
- **Maternal total caffeine consumption:** To calculate the total caffeine intake, mg/day of coffee and tea were summed up. For this analysis women with missing data on both coffee and tea were excluded. NAs were treated as 0, unless tea *and* coffee were missing, in which case the variable was treated as NA.
- **Maternal education** (as proxy for maternal socioeconomic position): Maternal education was assessed via a questionnaire sent out during pregnancy (after enrolment) and coded into three categories: ‘no education or primary education’, ‘secondary education/high school’, ‘higher education (college or university)’
- **Maternal smoking during pregnancy** was assessed via questionnaires, sent out during each trimester of pregnancy and coded as three categories: ‘no smoking in pregnancy/quit smoking until pregnancy was known (i.e., first trimester only)’ or ‘continued smoking throughout pregnancy’
- **Maternal age** continuous numeric variable in years assessed via a questionnaire sent out during pregnancy
- **Maternal BMI** was calculated from measured height during a visit to the research center during the first trimester of pregnancy and self-reported pre-pregnancy weight assessed via a questionnaire sent out during pregnancy
- **Parity** was assessed via a questionnaire sent out during pregnancy
- **Gestational age** was calculated (in days) based on the date of the mother’s last menstrual period (LMP) when the mother was certain of this and if she had a regular menstrual cycle of 28±4 days. For uncertain LMP and/or irregular cycle, ultrasound assessment was used.
- **Offspring sex** was recorded by midwife/hospital records.

K5 INfancia y Medio Ambiente (INMA) definition of variables and covariates

Dr. Silvia Fernandez who performed the analysis in INMA provided the description of variables used in INMA

- **Maternal coffee consumption:** Coffee intake was assessed in cups per day at week 12 of pregnancy. A semi-quantitative FFQ of 101 items was used with 9 possible responses from "never or less than once per month" to "six or more per day". Based on these responses daily consumption of coffee was calculated. Cups of coffee were then transformed to mg/day by: cups/day*57.
- **Maternal tea consumption:** Tea and herbal infusion intake was assessed in cups per day at week 12 of pregnancy in the same way as coffee consumption. Cups of tea were then transformed to mg/day assuming 27 mg of caffeine per cup.
- **Maternal cola consumption:** Regular and light soda intake was assessed in glasses per day at week 12 of pregnancy. Glasses of soda were then transformed to mg/day assuming 20mg of caffeine per glass of cola.
- **Maternal total caffeine consumption:** Total maternal caffeine intake during pregnancy in mg/day, summing caffeine from tea, coffee and cola drinks.
- **Maternal education** (as proxy for maternal socioeconomic position): Maternal education was assessed at week 12 of gestation and coded as an ordinal variable: 1 = without studies/primary studies unfinished, 2 = primary studies, 3 = secondary, and 4 = University.
- **Maternal smoking during pregnancy** was assessed as an ordinary variable representing 0 = no or quit smoking before second trimester, 2 = Smoking in the third trimester or throughout pregnancy.
- **Maternal age** continuous numeric variable in years assessed at enrolment.
- **Maternal BMI** Maternal pre-pregnancy BMI was calculated from measured height and self-reported pre-pregnancy weight collected using a questionnaire at enrolment (week 12 of pregnancy). Reported pre-pregnancy weight was highly correlated with measured weight at 12 weeks of pregnancy in INMA ($r = 0.96$; $P < 0.0001$).
- **Parity** based on previous born children (previous stillbirths included, abortions excluded), asked at 12 weeks assessment. It was coded as 0 = no previous pregnancy and 1 = one or more previous pregnancies.
- **Gestational age** was calculated (in weeks) based on the date of the mother's last menstrual period when the mother was certain of this, corrected with the information from the ultrasound assessment. It was transformed to days by multiplying it by 7.
- **Offspring sex** was taken from obstetric records.

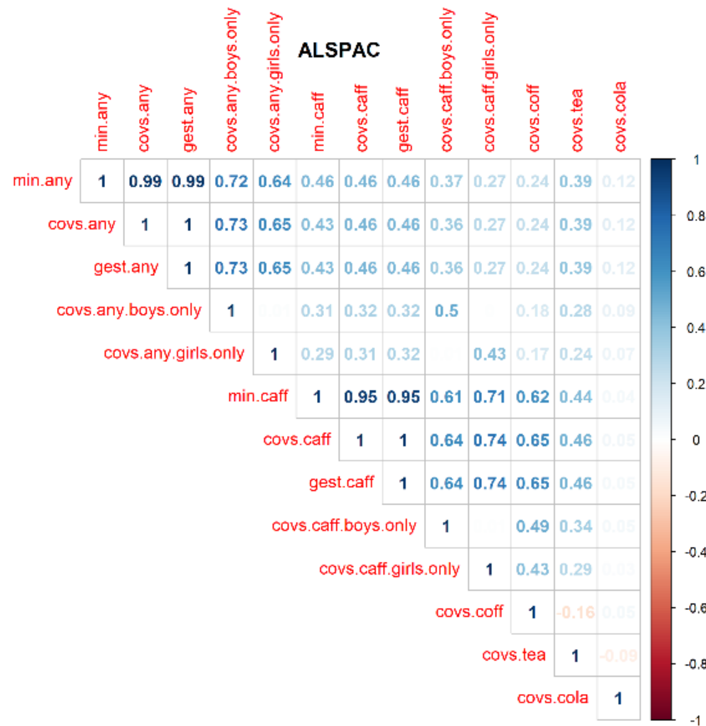
K6 EDEN definition of variables and covariates

Dr. Giancarlo Pesce who performed the analysis in EDEN provided the description of variables used in EDEN

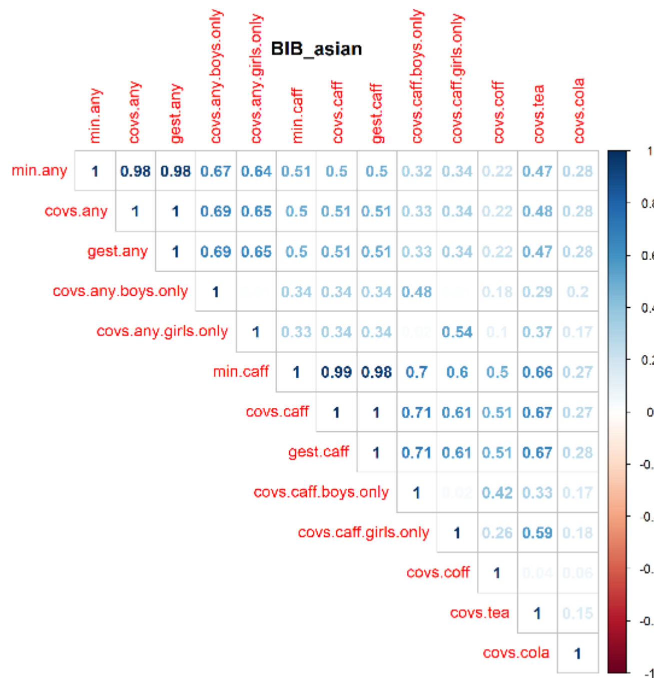
- **Maternal coffee consumption:** Coffee intake was initially assessed using a questionnaire administered during the week 24-28 of gestation, where mothers reported the daily consumption in the first trimester of pregnancy. Total cups of coffee were derived by adding up the cups of coffee consumed per day at home and outdoors. Cups of coffee were then transformed to mg/day by the formula: cups/week*57.
- **Maternal tea consumption:** Tea intake was initially assessed using a questionnaire administered during the week 24-28 of gestation, where mothers reported the daily consumption in the first trimester of pregnancy. Total cups of tea were derived by adding up the cups of tea consumed per day at home and outdoors. Cups of tea were then transformed to mg/day by the formula: cups/week*27.
- **Maternal cola consumption:** regular and light cola intakes were initially assessed using a food frequency questionnaire administered after delivery about the dietary habits in the last trimester of pregnancy. Total volume of cola consumption per day was estimated by multiplying the frequency of consumption by the volume of the most used glass for cola drinking. Total cups of cola were derived by adding up the volumes of regular and light cola consumption per day, divided by 250mL. Cups were then transformed to mg/day by the formula: cups*20.
- **Maternal total caffeine consumption:** Total maternal caffeine intake during pregnancy in mg/day, summing caffeine from tea, coffee and cola drinks. NAs were treated as 0, unless tea, coffee *and* cola were all missing, in which case the variable was coded as missing.
- **Maternal education** (as proxy for maternal socioeconomic position): Highest degree obtained by the mother was assessed at week 24-28 of gestation, as self-reported by the mother, and coded as an ordinal variable: "Vocational school" = 0, “French *baccalauréat* (BAC) degree” = 1, "BAC + 2 years of College" = 2, and “University or higher degree” = 3.
- **Maternal smoking during pregnancy** was assessed as an ordinary variable representing 0 = no or early smoking during pregnancy, 1 = Smoking throughout pregnancy.
- **Maternal age** continuous numeric variable in years assessed at birth of study child
- **Maternal BMI** was assessed continuous numeric variable in kg/m^2, dividing the weight before the beginning of the pregnancy, as self-reported by the mother at the clinical interview during pregnancy, and squared height in meters, as measured in clinic during pregnancy.
- **Parity** coded as 0 = nulliparous, 1 = with at least one living child.
- **Gestational age** was included as continuous variable (in weeks) based on the date of the mother’s last menstrual period (LMP) and the date of birth of the child.
- **Offspring sex** was taken from obstetric records.

Appendix L

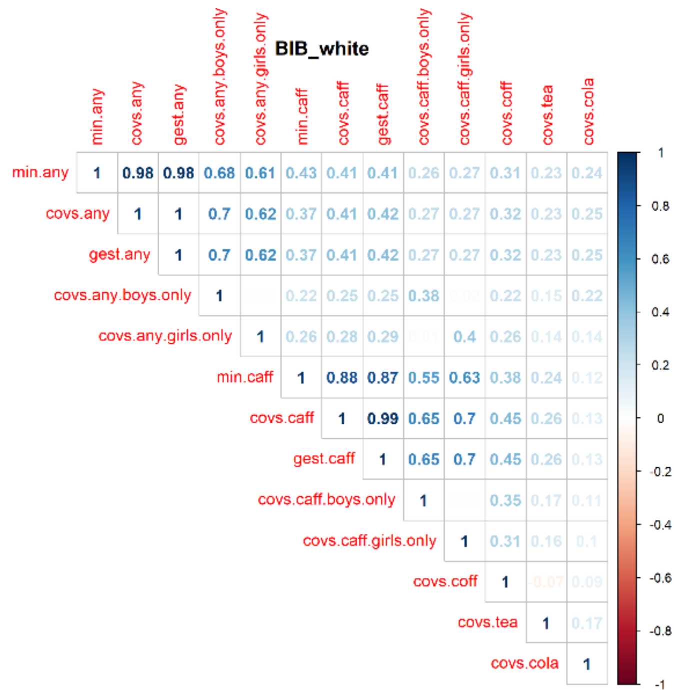
L1 Correlation plot ALSPAC



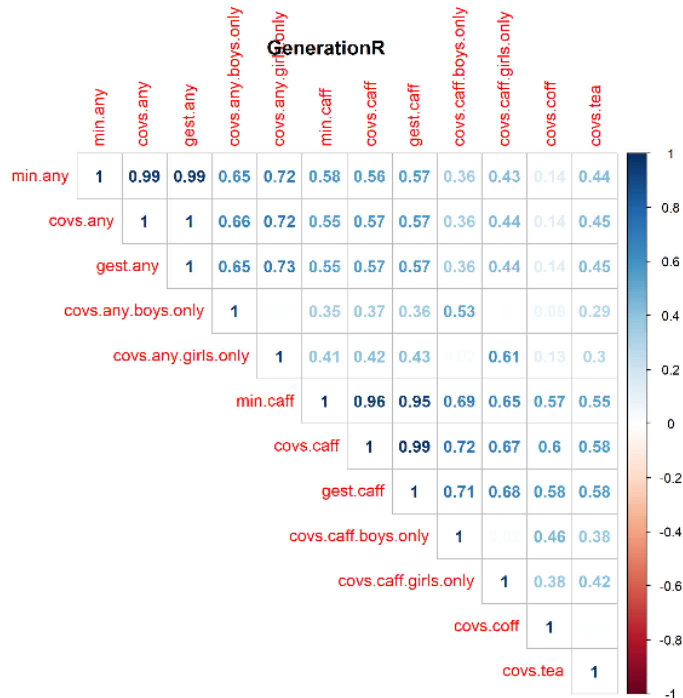
L2 Correlation plot BiB (Asian ethnicity)



L3 Correlation plot BiB (White European ethnicity)

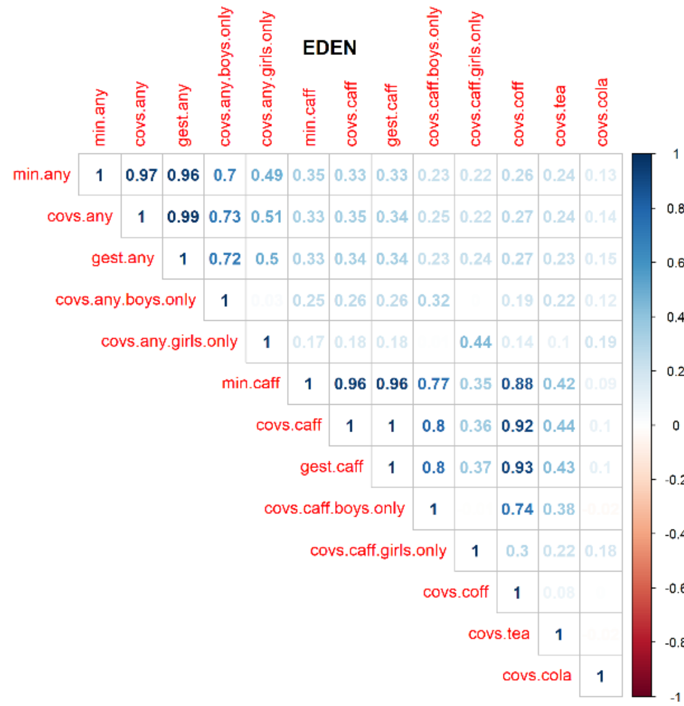


L4 Correlation plot Generation R

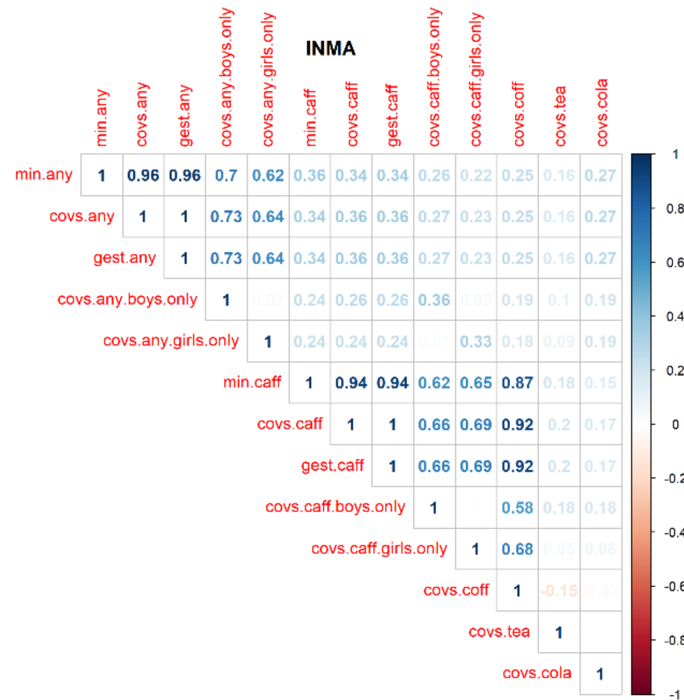


Correlation plot EDEN

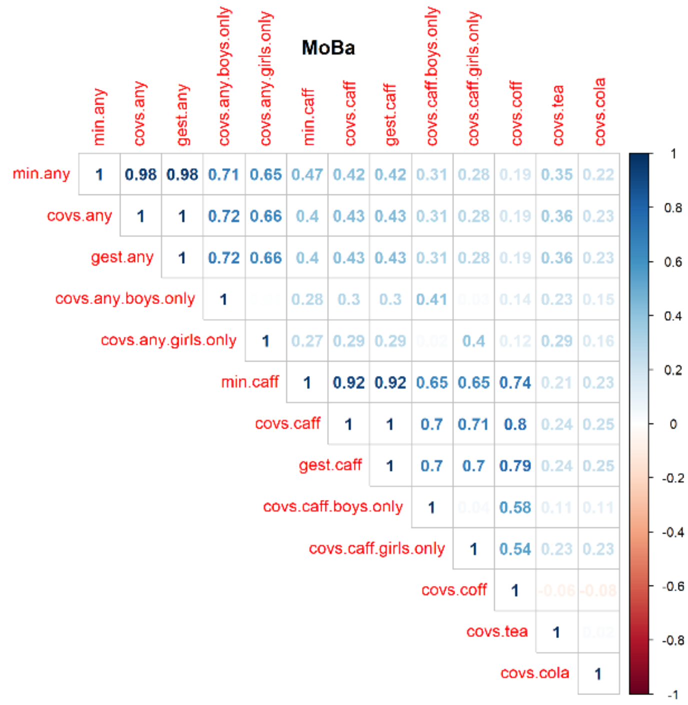
Figure P1



L5 Correlation plot INMA

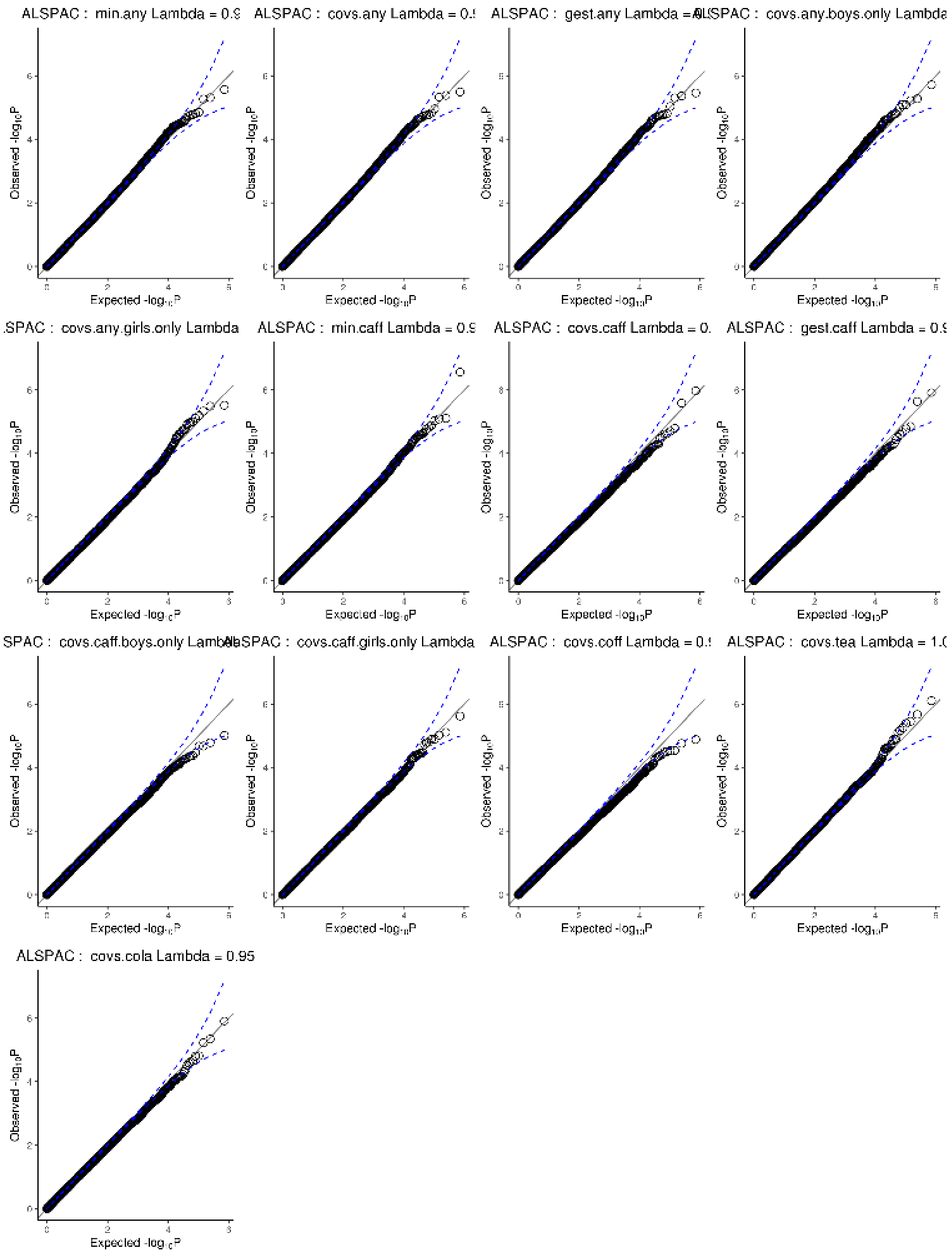


L6 Correlation plot MoBa1

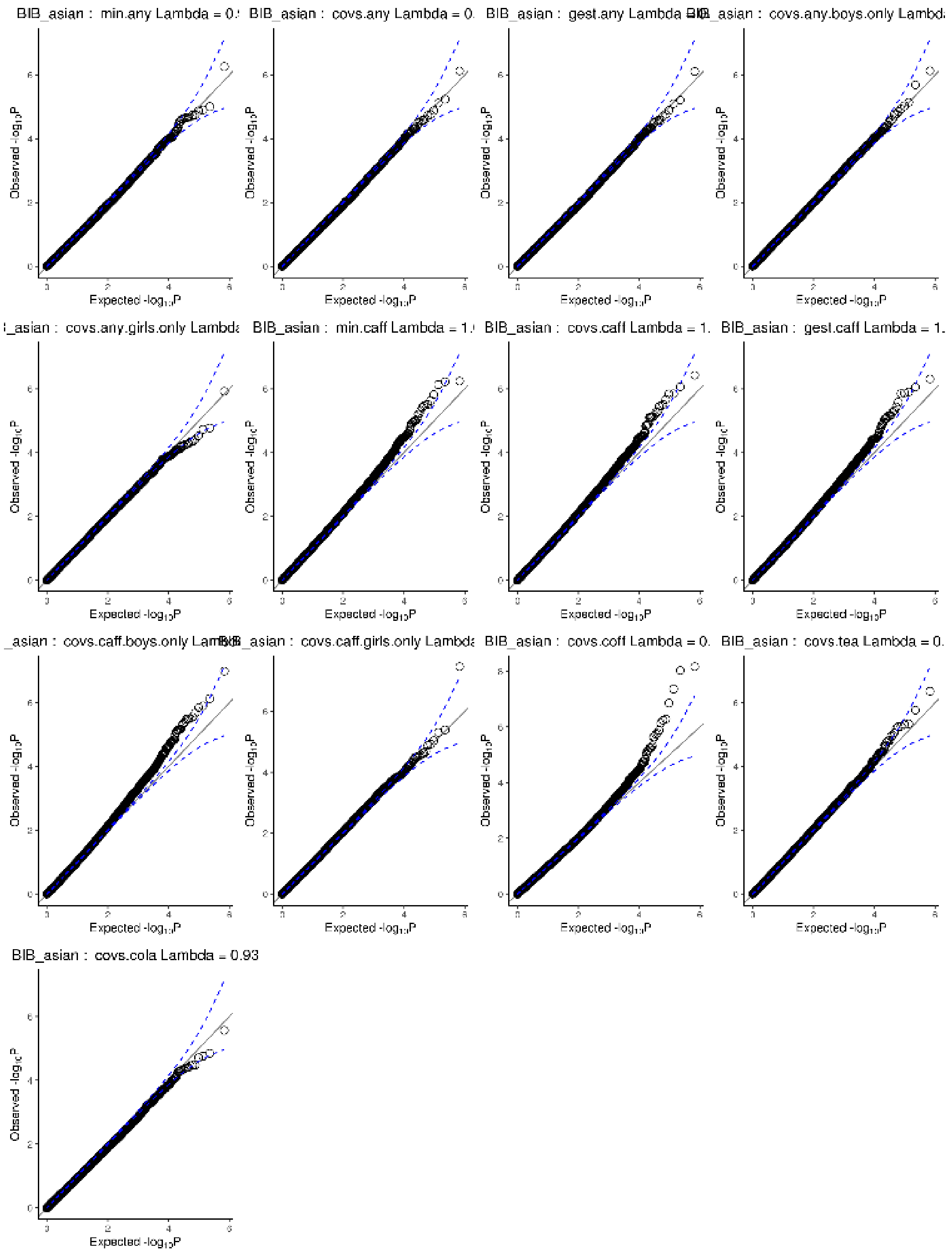


Appendix M

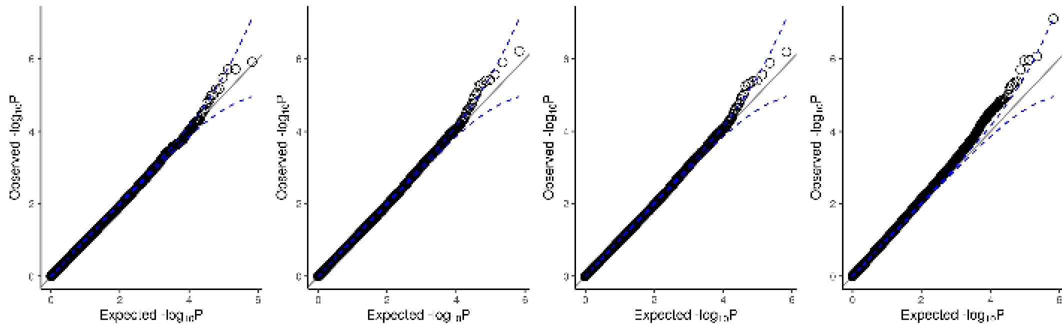
M1 ALSPAC QQ-Plots of caffeine models

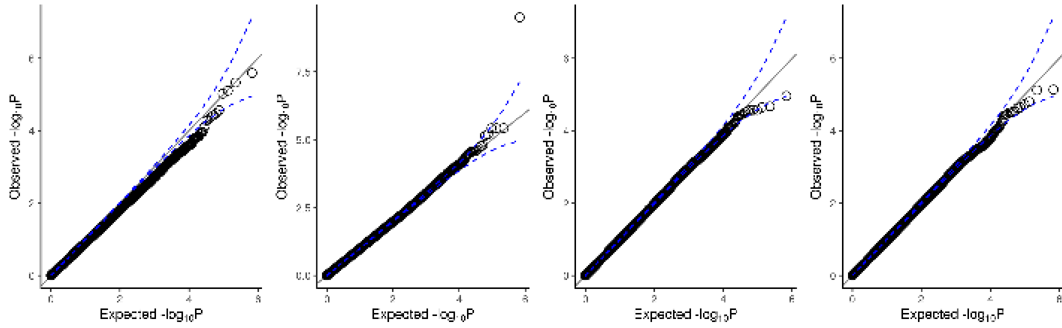


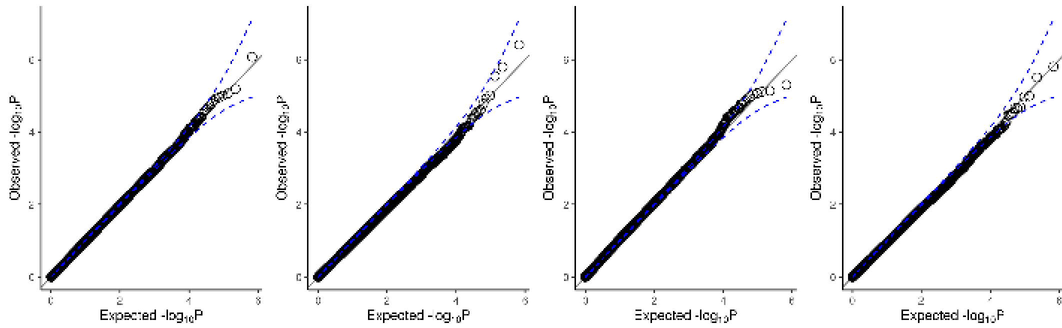
M2 BiB (Asian ethnicity) QQ-Plots of caffeine models

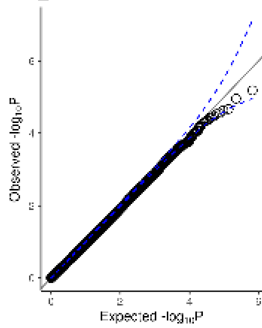


M3 BiB (White European ethnicity) QQ-Plots of caffeine models

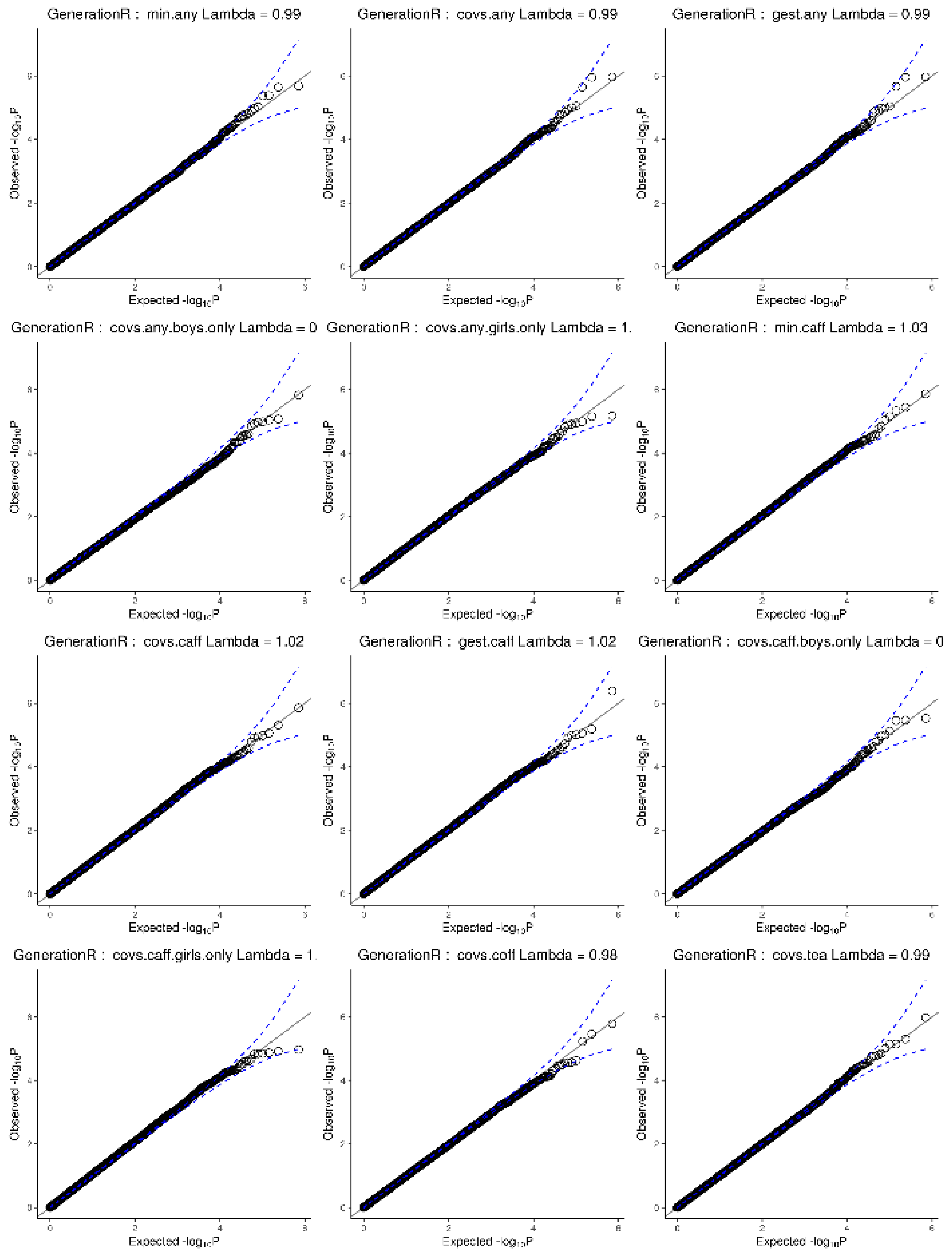




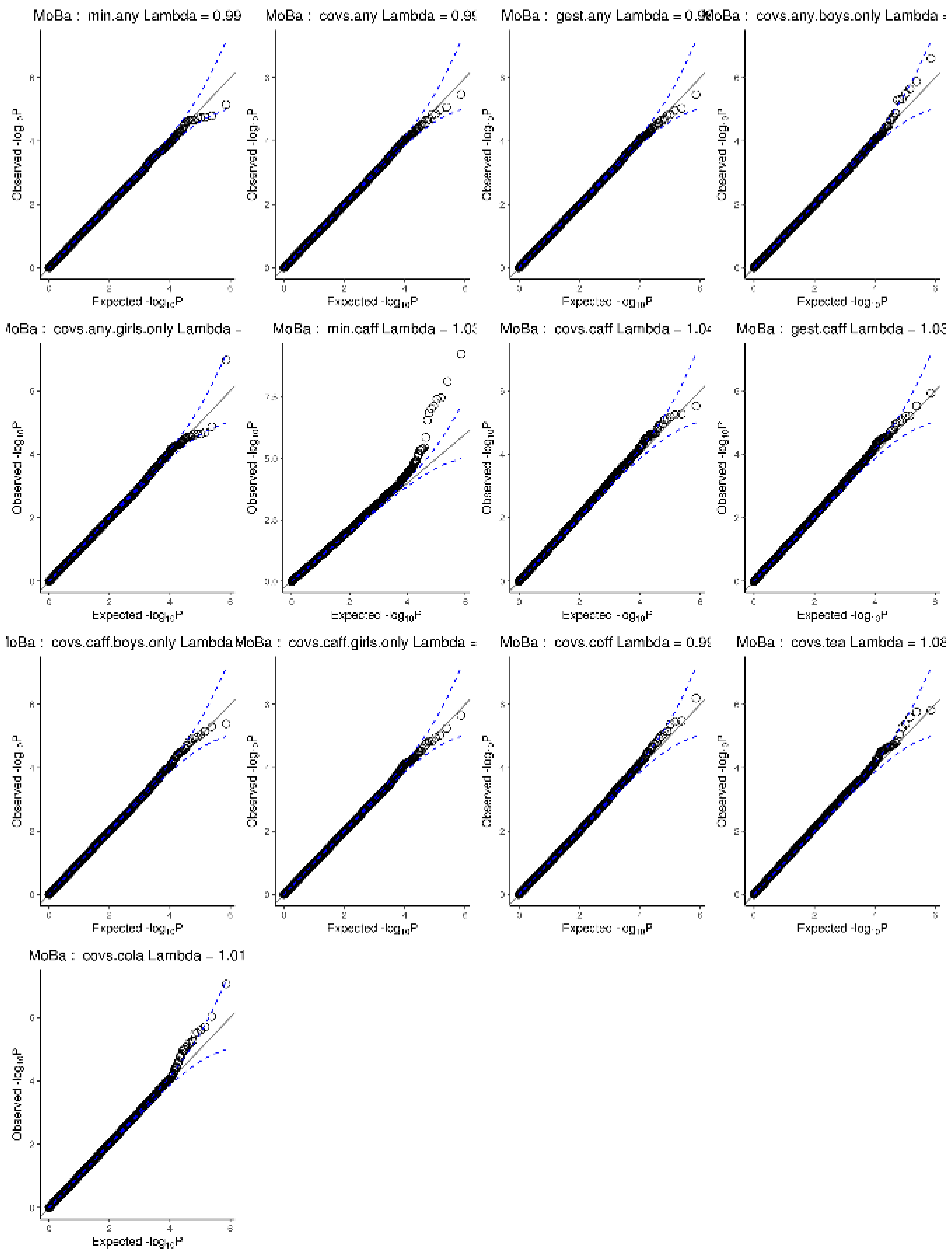




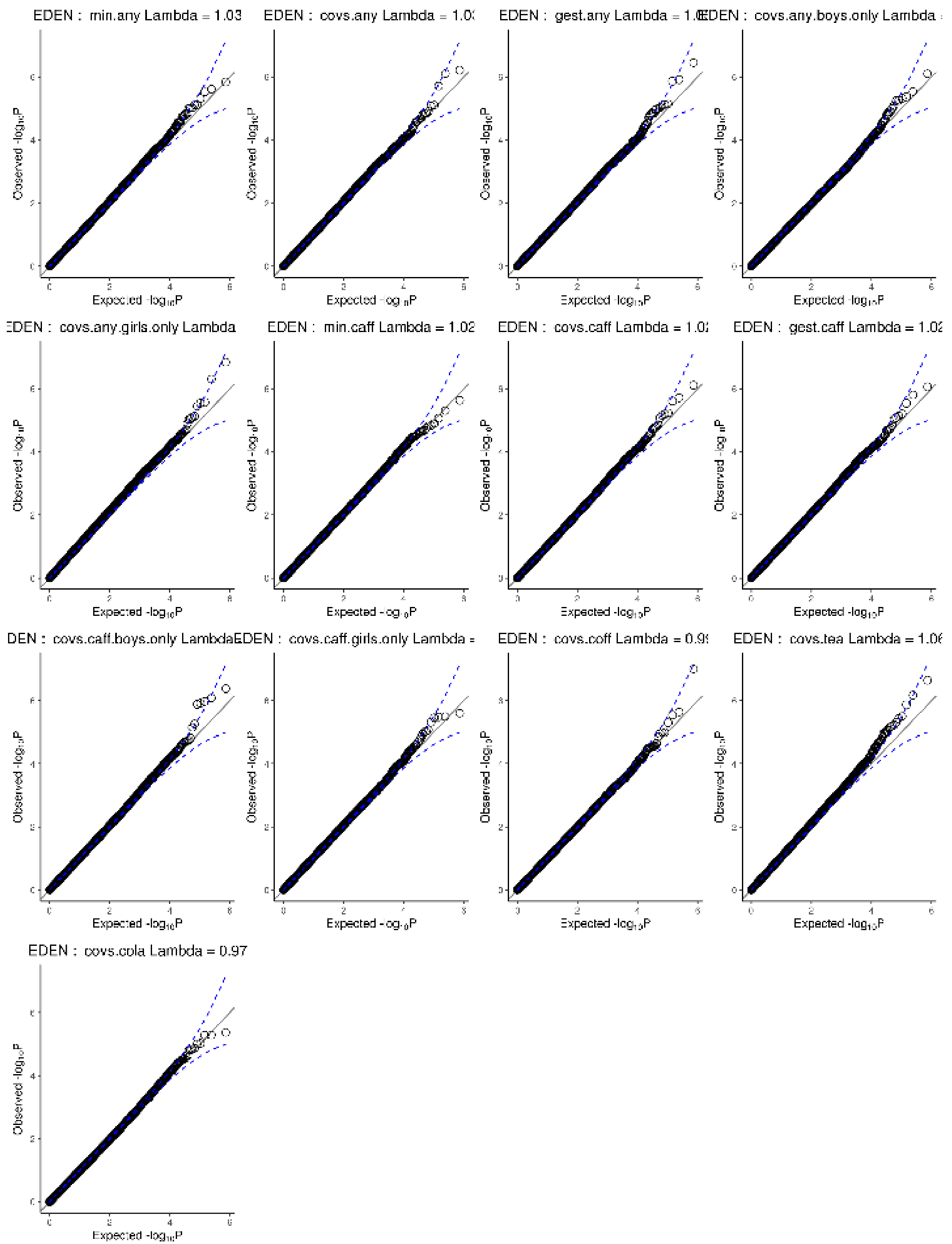
M4 Generation R QQ-Plots of caffeine models



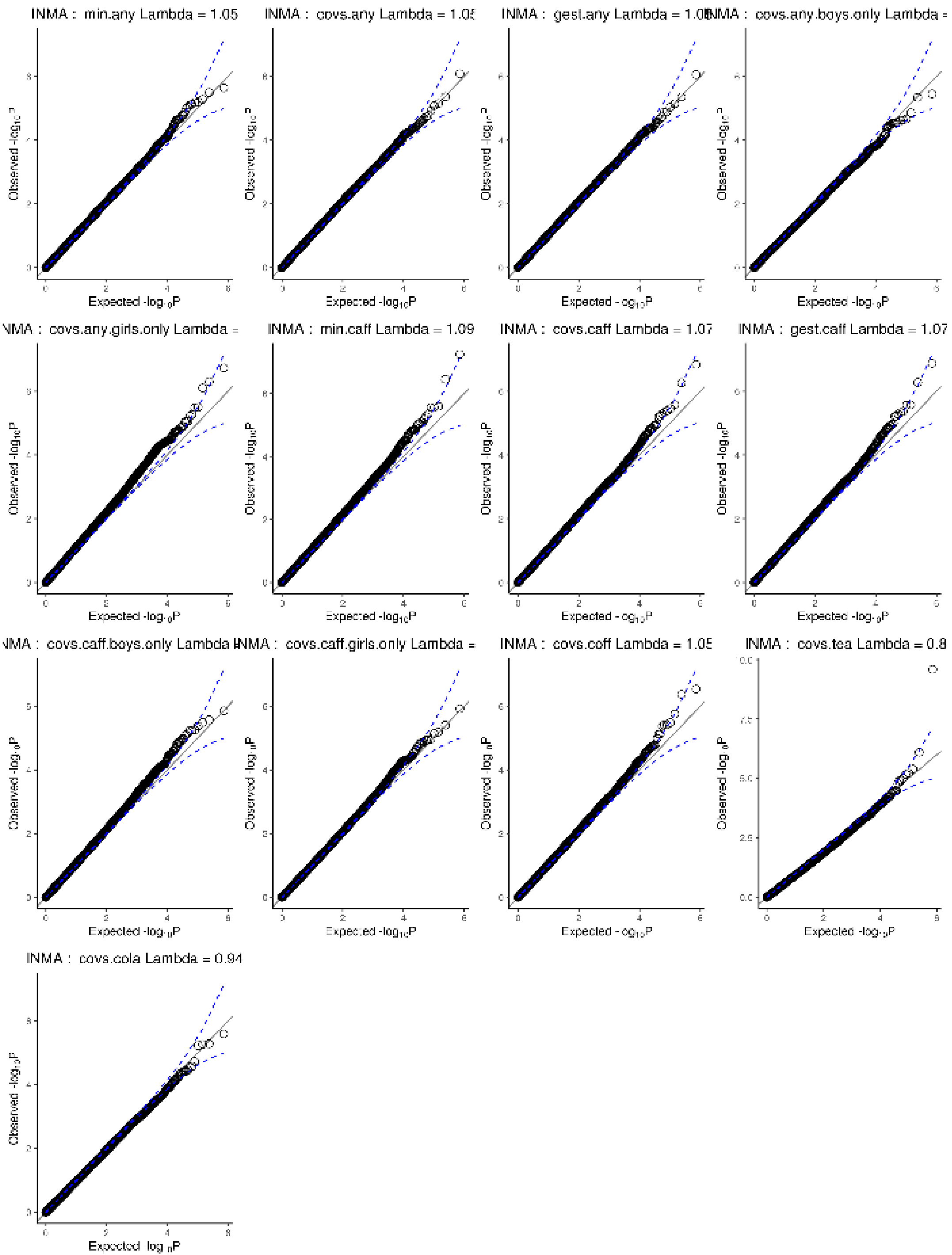
M5 MoBa1 QQ-Plots of caffeine models



M6 EDEN QQ-Plots of caffeine models

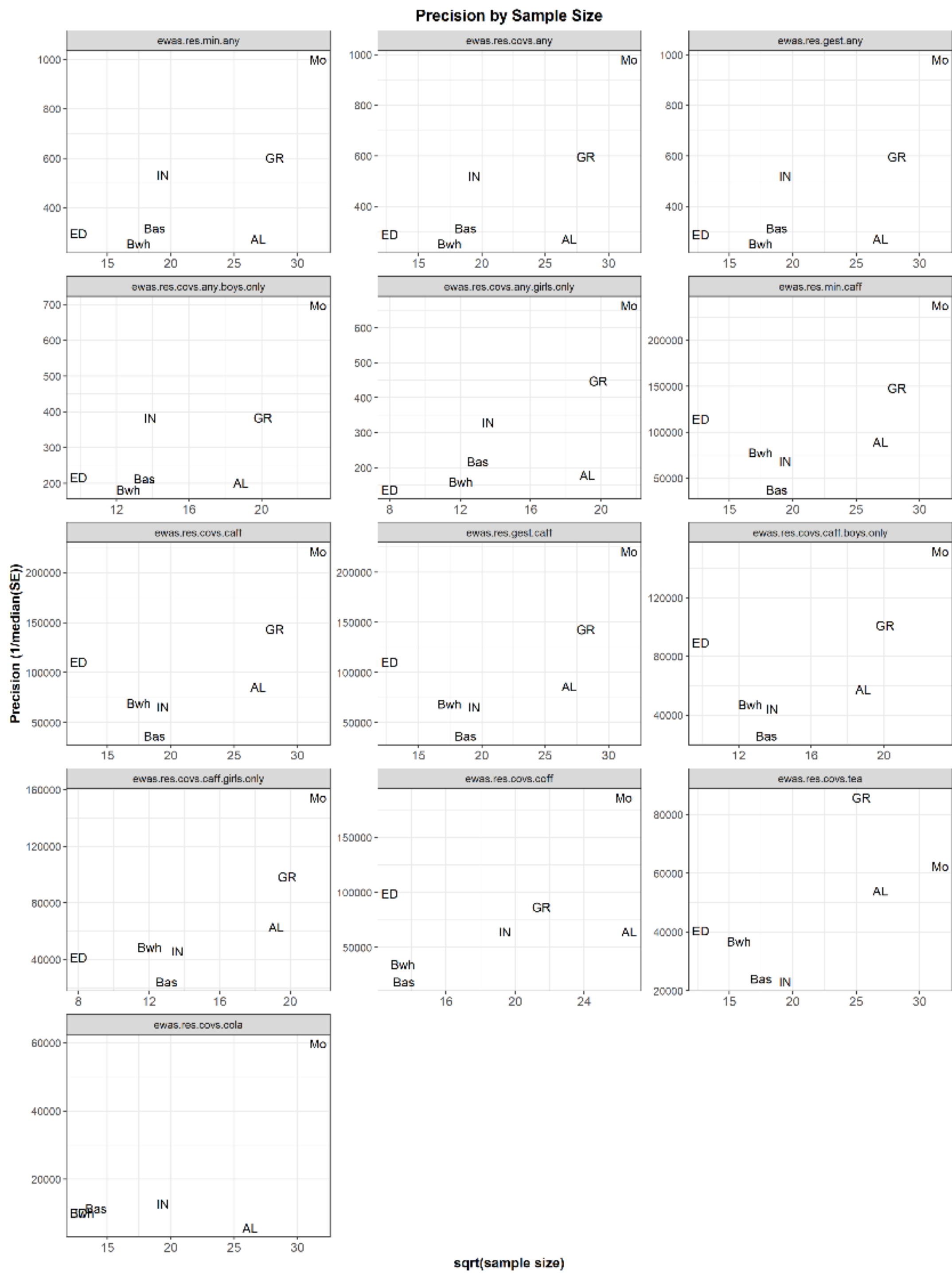


M7 INMA QQ-Plots of caffeine models



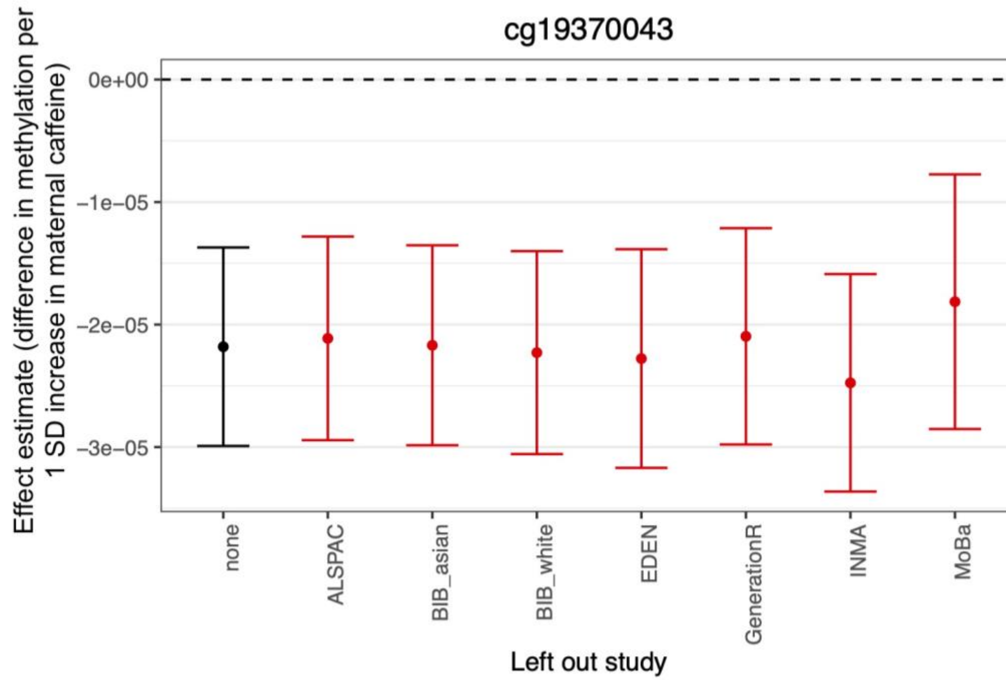
Appendix N

N1 Precision plots

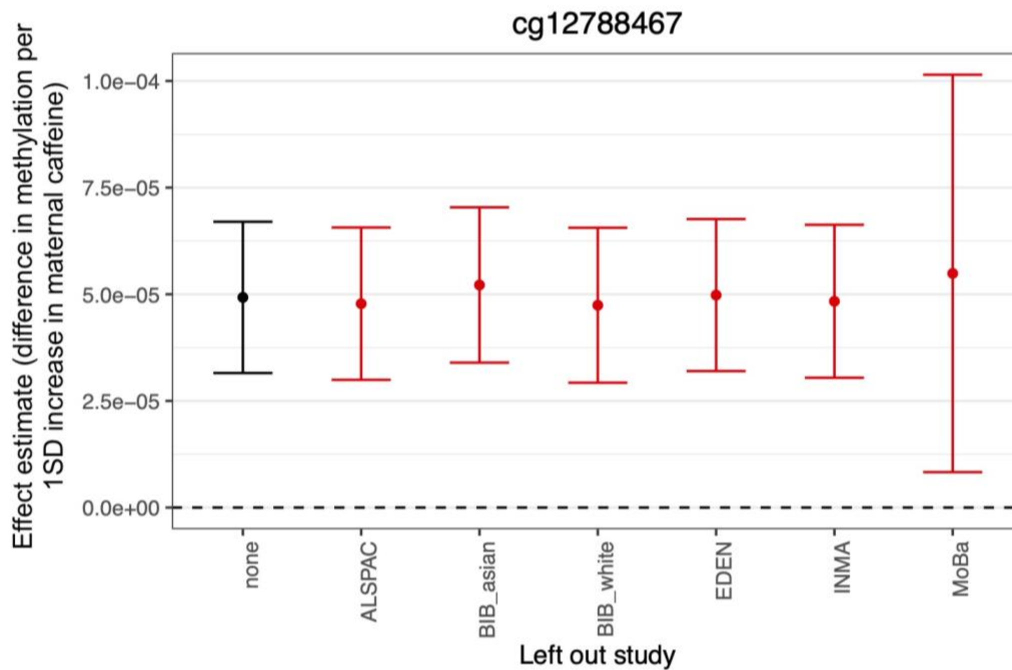


Appendix O

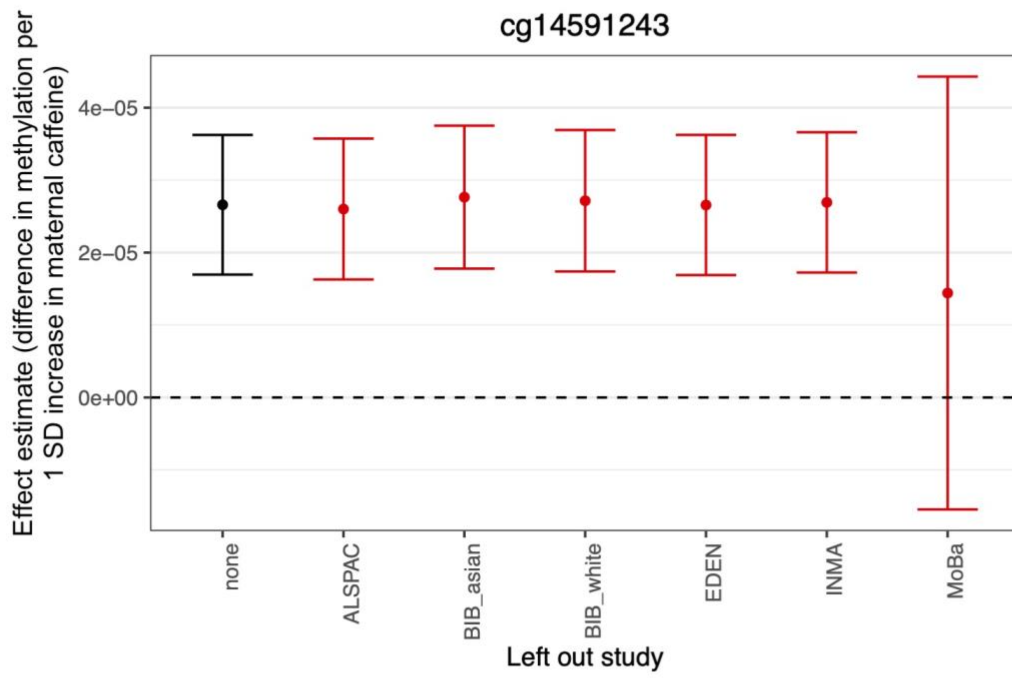
O1 Leave-one-out plot of the prenatal caffeine-associated CpG site (Cg19370043)



O2 Leave-one-out plot of the prenatal cola-associated CpG site (Cg12788467)

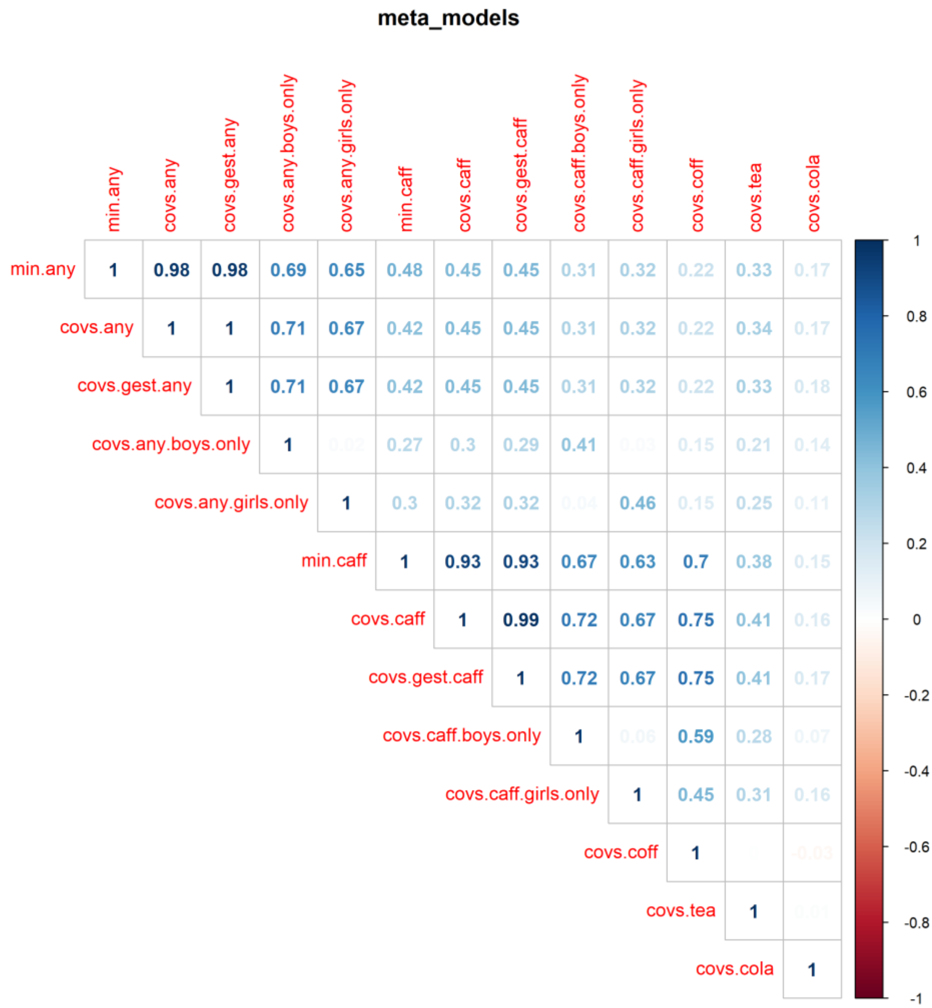


O3 Leave-one-out plot of the prenatal cola-associated CpG site (Cg14591243)



Appendix P

P1 Correlation matrix of the meta-analysed caffeine models



Appendix Q

Q1 Top five associations between maternal caffeine polygenic risk score (PRS) and offspring DNA methylation

CpG (gene)	**Estimate (SE)**	**P-value**
cg23254346 (ELAC2)	-0.55 (0.12)	5.63×10^{-06}
cg03136668	-1.34 (0.27)	3.51×10^{-06}
cg10414208	1.01 (0.24)	3.99×10^{-05}
cg00157109 (PRKAG2)	0.78 (0.19)	3.87×10^{-05}
cg02816367 (FYN)	-0.82(0.19)	1.45×10^{-05}

Note. SE = standard error.

Appendix R

R1 Results from the meta-analysis of differentially methylated regions for maternal total caffeine consumption

Differentially-methylated region (DMR)	N consecutive CpG sites	Estimate (SE)	Z	P-value	Bonferroni adjusted P-value	Gene
chr17:58499679-58499911	7	-3.77×10^{-05} (5.02×10^{-06})	-8.15	3.650×10^{-16}	1.420×10^{-10}	C17orf64
chr11:6291339-6292490	9	-5.57×10^{-05} (7.34×10^{-06})	-7.59	3.270×10^{-14}	1.270×10^{-08}	CCKBR
chr12:47219737-47220092	10	-3.77×10^{-05} (5.02×10^{-06})	-7.52	5.680×10^{-14}	2.210×10^{-08}	SLC38A4
chr6:30094980-30095341	14	-3.47×10^{-05} (4.86×10^{-06})	-7.14	9.560×10^{-13}	3.720×10^{-07}	-
chr6:33245488-33245770	15	-1.57×10^{-05} (2.25×10^{-06})	-6.98	2.850×10^{-12}	1.110×10^{-06}	B3GALT4
chr20:61446962-61447369	11	2.57×10^{-05} (4.12×10^{-06})	6.24	4.500×10^{-10}	1.753×10^{-04}	COL9A3
chr10:63657059-63657363	3	6.59×10^{-05} (1.08×10^{-05})	6.10	1.040×10^{-09}	4.053×10^{-04}	-
chr6:29599160-29599331	8	4.07×10^{-05} (6.70×10^{-06})	6.08	1.230×10^{-09}	4.782×10^{-04}	GABBR1
chr5:140729653-140730516	7	-2.96×10^{-05} (5.01×10^{-06})	-5.91	3.340×10^{-09}	1.301×10^{-03}	PCDHGA2;PCDHGB1;PCDHGA1;PCDHGA3
chr1:117317903-117318185	4	3.51×10^{-05} (5.94×10^{-06})	5.90	3.540×10^{-09}	1.380×10^{-03}	-
chr22:38713874-38714416	8	-2.14×10^{-05} (3.77×10^{-06})	-5.68	1.360×10^{-08}	5.310×10^{-03}	CSNK1E
chr7:130130588-130131258	12	1.60×10^{-05} (2.83×10^{-06})	5.65	1.600×10^{-08}	6.223×10^{-03}	MESTIT1;MEST
chr6:41410759-41411128	4	-1.33×10^{-05} (2.40×10^{-06})	-5.56	2.660×10^{-08}	1.035×10^{-02}	-
chr7:27153580-27153847	6	-3.66×10^{-05} (6.60×10^{-06})	-5.55	2.810×10^{-08}	1.094×10^{-02}	HOXA3
chr10:50649723-50650248	4	4.89×10^{-05} (8.84×10^{-06})	5.54	3.100×10^{-08}	1.209×10^{-02}	-
chr6:149806131-149806339	4	-3.18×10^{-05} (5.74×10^{-06})	-5.53	3.200×10^{-08}	1.245×10^{-02}	ZC3H12D
chr6:32164723-32165237	9	-2.46×10^{-05} (4.47×10^{-06})	-5.51	3.640×10^{-08}	1.416×10^{-02}	GPSM3;NOTCH4
chr4:62383028-62383240	3	-4.34×10^{-05} (7.95×10^{-06})	-5.46	4.790×10^{-08}	1.867×10^{-02}	LPHN3
chr1:228400217-228400419	2	4.11×10^{-05} (7.66×10^{-06})	5.37	8.020^{-08}	3.125×10^{-02}	OBSCN
chr6:7468673-7468973	4	-3.68×10^{-05} (6.86×10^{-06})	-5.36	8.340^{-08}	3.247×10^{-02}	-
chr11:368366-368943	15	-2.03×10^{-05} (3.80×10^{-06})	-5.34	9.250^{-08}	3.601×10^{-02}	B4GALNT4
chr7:27142810-27143403	8	2.96×10^{-05} (5.57×10^{-06})	5.32	1.060^{-07}	4.145×10^{-02}	HOXA2

R2 Results from the meta-analysis of differentially methylated regions for any vs. no maternal caffeine consumption.

Differentially-methylated region (DMR)	N consecutive CpG sites	Estimate (SE)	Z	P-value	Bonferroni adjusted P-value	Gene
chr6:31734147-31734554	10	9.44×10^{-03} (1.37×10^{-03})	6.91	4.975×10^{-12}	1.928×10^{-06}	C6orf27
chr1:11561497-11562134	5	1.05×10^{-02} (1.63×10^{-03})	6.42	1.338×10^{-10}	5.184×10^{-05}	PTCHD2
chr22:51016501-51017162	8	1.16×10^{-02} (1.83×10^{-03})	6.32	2.691×10^{-10}	1.043×10^{-04}	CPT1B; CHKB-CPT1B
chr17:39969264-39969297	3	-5.99×10^{-03} (9.97×10^{-04})	-6.01	1.894×10^{-09}	7.339×10^{-04}	SC65;FKBP10
chr17:19883326-19883474	2	-1.24×10^{-02} (2.13×10^{-03})	-5.80	6.495×10^{-09}	2.517×10^{-03}	-
chr20:61002657-61002866	5	8.23×10^{-03} (1.43^{-03})	5.75	8.997×10^{-09}	3.487×10^{-03}	C20orf151
chr11:64980819-64981297	5	8.00×10^{-03} (1.44×10^{-03})	5.57	2.553×10^{-08}	9.893×10^{-03}	SLC22A20
chr10:530714-531152	6	1.08×10^{-02} (1.94×10^{-03})	5.55	2.885×10^{-08}	1.118×10^{-02}	DIP2C
chr17:40575479-40575822	6	3.15×10^{-03} (5.70×10^{-04})	5.53	3.178×10^{-08}	1.232×10^{-02}	PTRF
chr9:136567339-136568145	3	9.21×10^{-03} (1.73×10^{-03})	5.34	9.508×10^{-08}	3.685×10^{-02}	SARDH
chr20:61447181-61447490	10	4.58×10^{-03} (8.63×10^{-04})	5.31	1.080×10^{-07}	4.186×10^{-02}	COL9A3

R3 Table H3 Results from the meta-analysis of differentially methylated regions for maternal coffee consumption.

Differentially-methylated region (DMR)	N consecutive CpG sites	Estimate (SE)	Z	P-value	Bonferroni adjusted P-value	Gene
chr1:25291385-25292034	7	-4.59×10^{-05} (5.87×10^{-06})	-7.83	4.831×10^{-15}	1.877×10^{-09}	RUNX3
chr22:46449461-46449821	5	-9.14×10^{-05} (1.26×10^{-05})	-7.26	3.853×10^{-13}	1.497×10^{-07}	C22orf26; LOC150381
chr11:2019732-2020314	19	-2.22×10^{-05} (3.33×10^{-06})	-6.67	2.528×10^{-11}	9.824×10^{-06}	H19
chr6:49681178-49681742	8	-9.79×10^{-05} (1.55×10^{-05})	-6.30	2.911×10^{-10}	1.131×10^{-04}	CRISP2
chr11:2292895-2293173	10	6.05×10^{-05} (1.01×10^{-05})	5.98	2.212×10^{-09}	8.594×10^{-04}	ASCL2
chr1:63249197-63249213	4	8.92×10^{-05}(1.60×10^{-05})	5.57	2.537×10^{-08}	9.859×10^{-03}	ATG4C
chr7:5111621-5111916	5	5.33×10^{-05} (9.61×10^{-06})	5.55	2.851×10^{-08}	1.108×10^{-02}	LOC389458
chr11:368351-368683	11	-4.43×10^{-05}($8.10^{-} \times 10^{06}$)	-5.47	4.496×10^{-08}	1.747×10^{-02}	B4GALNT4
chr17:58499679-58499911	7	-3.93×10^{-05}(7.20×10^{-06})	-5.46	4.866×10^{-08}	1.891×10^{-02}	C17orf64
chr6:155537901-155538055	5	-4.43×10^{-05}(8.20×10^{-06})	-5.40	6.837×10^{-08}	2.657×10^{-02}	TIAM2
chr2:237478526-237478664	3	-4.98×10^{-05}(9.24×10^{-06})	-5.39	6.913×10^{-08}	2.686×10^{-02}	CXCR7
chr2:732577-732961	2	-1.66×10^{-04}(3.13×10^{-05})	-5.30	1.169×10^{-07}	4.542×10^{-02}	-

R4 Results from the meta-analysis of differentially methylated regions for maternal tea consumption.

Differentially-methylated region (DMR)	N consecutive CpG sites	Estimate (SE)	Z	P-value	Bonferroni adjusted P-value	Gene
chr13:114814024-114814401	4	-1.06×10^{-04} (1.54×10^{-05})	-6.88	6.181×10^{-12}	2.393×10^{-06}	RASA3
chr7:27143046-27143370	6	6.86×10^{-05} (1.04×10^{-05})	6.59	4.355×10^{-11}	1.686×10^{-05}	HOXA2
chr6:31127120-31127379	7	5.47×10^{-05} (8.66×10^{-06})	6.31	2.766×10^{-10}	1.071×10^{-04}	TCF19; CCHCR1
chr6:160023581-160024002	4	1.32×10^{-04} (2.25×10^{-05})	5.88	4.198×10^{-09}	1.625×10^{-03}	
chr6:29599160-29599331	8	6.55×10^{-05} (1.12×10^{-05})	5.82	5.802×10^{-09}	2.247×10^{-03}	GABBR1
chr11:2019862-2020537	18	2.10×10^{-05} (3.61×10^{-06})	5.81	6.068×10^{-09}	2.350×10^{-03}	H19
chr19:51189671-51190179	4	7.05×10^{-05} (1.22×10^{-05})	5.80	6.744×10^{-09}	2.611×10^{-03}	SHANK1
chr3:42705828-42706106	3	-6.61×10^{-05} (1.15×10^{-05})	-5.73	1.017×10^{-08}	3.940×10^{-03}	ZBTB47
chr6:136872115-136872119	2	8.55×10^{-05} (1.49×10^{-05})	5.73	1.021×10^{-08}	3.955×10^{-03}	MAP7
chr1:117317838-117318185	5	4.62×10^{-05} (8.07×10^{-06})	5.73	1.031×10^{-08}	3.992×10^{-03}	-
chr7:155284062-155284759	5	8.80×10^{-05} (1.55×10^{-05})	5.69	1.242×10^{-08}	4.811×10^{-03}	-
chr8:144659883-144660772	5	5.20^{-05} (9.16^{-06})	5.67	1.422×10^{-08}	5.507×10^{-03}	NAPRT1
chr6:32015651-32015737	3	1.94×10^{-05} (3.51×10^{-06})	5.53	3.217×10^{-08}	1.246×10^{-02}	TNXB
chr9:136198682-136199172	3	-1.91×10^{-05} (3.52×10^{-06})	-5.41	6.305×10^{-08}	2.442×10^{-02}	SURF6
chr10:63657059-63657363	3	9.72×10^{-05} (1.81×10^{-05})	5.37	7.766×10^{-08}	3.007×10^{-02}	
chr17:80976690-80977389	3	2.72×10^{-05} (5.12×10^{-06})	5.32	1.033×10^{-07}	4.001×10^{-02}	B3GNTL1
chr1:234367145-234367493	4	1.14×10^{-04} (2.15×10^{-05})	5.32	1.051×10^{-07}	4.071×10^{-02}	SLC35F3
chr17:58499679-58499816	5	-4.68×10^{-05} (8.86×10^{-06})	-5.28	1.282×10^{-07}	4.966×10^{-02}	C17orf64

R5 Results from the meta-analysis of differentially methylated regions for maternal cola consumption.

Differentially-methylated region (DMR)	N consecutive CpG sites	Estimate (SE)	Z	P-value	Bonferroni adjusted P-value	Gene
chr3:99979117-99979421	3	-3.48×10^{-04} (4.06×10^{-05})	-8.57	1.024×10^{-17}	4.087×10^{-12}	TBC1D23
chr3:37033632-37033980	5	-1.49×10^{-03} (2.06×10^{-04})	-7.25	4.099×10^{-13}	1.636×10^{-07}	EPM2AIP1;MLH1
chr22:31318103-31318546	9	-4.96×10^{-03} (6.99×10^{-04})	-7.10	1.218×10^{-12}	4.862×10^{-07}	C22orf27
chr2:113992762-113993313	7	-5.24×10^{-03} (7.70×10^{-04})	-6.80	1.032×10^{-11}	4.118×10^{-06}	PAX8
chr1:242220538-242220925	3	-5.53×10^{-03} (8.52×10^{-04})	-6.49	8.403×10^{-11}	3.353×10^{-05}	
chr2:11101549-11101592	2	-8.33×10^{-03} (1.33×10^{-03})	-6.27	3.533×10^{-10}	1.410×10^{-04}	
chr7:30635762-30636176	4	-4.05×10^{-03} (6.56×10^{-04})	-6.18	6.499×10^{-10}	2.593×10^{-04}	GARS
chr14:50088544-50088598	2	-4.96×10^{-03} (8.06×10^{-04})	-6.15	7.553×10^{-10}	3.014×10^{-04}	RPL36AL;MGAT2
chr7:128530800-128531165	2	3.21×10^{-03} (5.37×10^{-04})	5.97	2.332×10^{-09}	9.307×10^{-04}	KCP
chr17:18761852-18761932	2	2.80×10^{-04} (5.05×10^{-05})	5.56	2.753×10^{-08}	1.099×10^{-02}	PRPSAP2
chr11:66104174-66104485	4	1.99×10^{-03} (3.59×10^{-04})	5.55	2.851×10^{-08}	1.138×10^{-02}	RIN1
chr4:122853963-122854405	6	5.01×10^{-03} (9.20×10^{-04})	5.45	5.099×10^{-08}	2.035×10^{-02}	TRPC3
chr13:24519920-24520348	3	-1.23×10^{-02} (2.25×10^{-03})	-5.45	5.124×10^{-08}	2.045×10^{-02}	
chr1:152080471-152081002	3	2.43×10^{-03} (4.50×10^{-04})	5.41	6.461×10^{-08}	2.578×10^{-02}	TCHH

R6 Results from the meta-analysis of differentially methylated regions for maternal caffeine consumption stratified by female sex

Differentially-methylated region (DMR)	N consecutive CpG sites	Estimate (SE)	Z	P-value	Bonferroni adjusted P-value	Gene
chr2:139538222-139539001	5	2.50×10^{-03} (3.44×10^{-04})	7.28	3.266×10^{-13}	1.305×10^{-07}	NXPH2
chr19:50931432-50931622	4	2.08×10^{-03} (3.26×10^{-04})	6.38	1.812×10^{-10}	7.240×10^{05}	SPIB
chr21:35831996-35832180	4	1.87×10^{-03} (3.05×10^{-04})	6.14	8.366×10^{-10}	3.343×10^{-04}	KCNE1
chr4:184930940-184931143	4	1.32×10^{-04} (2.25×10^{-05})	5.85	5.048×10^{-09}	2.017×10^{-03}	STOX2
chr6:33257050-33257788	18	5.78×10^{-05} (1.01×10^{-05})	5.74	9.618×10^{-09}	3.844×10^{-03}	WDR46
chr11:45715517-45715523	2	-7.31×10^{-04} (1.29×10^{-04})	-5.65	1.627×10^{-08}	6.501×10^{-03}	
chr6:168197699-168197921	3	6.31×10^{-04} (1.13×10^{-04})	5.59	2.208×10^{-08}	8.823×10^{-03}	C6orf123
chr6:27860893-27860984	3	2.37×10^{-04} (4.34×10^{-05})	5.48	4.322×10^{-08}	1.727×10^{-02}	HIST1H2AMHIST1H2BO
chr4:89618637-89618667	2	1.58×10^{-03} (2.90×10^{-04})	5.44	5.467×10^{-08}	2.185×10^{-02}	NAP1L5; HERC3
chr20:25565460-25565667	2	-8.45×10^{-04} (1.56×10^{-04})	-5.41	6.327×10^{-08}	2.528×10^{-02}	NINL
chr11:368588-368847	6	-5.26×10^{-04} (9.80×10^{-05})	-5.37	7.859×10^{-08}	3.141×10^{-02}	B4GALNT4
chr4:146018582-146018691	2	-1.08×10^{-03} (2.02×10^{-04})	-5.37	8.036×10^{-08}	3.212×10^{-02}	ANAPC10; ABCE1

R7 Results from the meta-analysis of differentially methylated regions for maternal caffeine consumption stratified by male sex

Differentially-methylated region (DMR)	N consecutive CpG sites	Estimate (SE)	Z	P-value	Bonferroni adjusted P-value	Gene
chr12:58132093-58132558	3	-1.90 x 10^{-03} (2.65 x 10^{-04})	-7.18	7.022 x 10^{-13}	2.837 x 10^{-07}	AGAP2
chr3:49170599-49171051	6	-6.61 x 10^{-04} (9.84 x 10^{-05})	-6.71	1.881 x 10^{-11}	7.598 x 10^{-06}	LAMB2
chr7:94286343-94286760	16	5.06^{-} x 10^{04} (7.56 x 10^{-05})	6.69	2.190 x 10^{-11}	8.847 x10^{-06}	SGCE; PEG10
chr13:36871943-36872346	9	1.01 x 10^{-04} (1.56 x 10^{-05})	6.44	1.219 x 10^{-10}	4.92 x 10^{-05}	C13orf38
chr2:3699195-3699563	5	5.21 x 10^{-04} (8.65 x 10^{-05})	6.02	1.697 x 10^{-09}	6.856 x 10^{-04}	-
chr8:95961618-95962383	6	-4.29 x 10^{-04} (7.24 x 10^{-05})	-5.93	3.047 x 10^{-09}	1.231 x 10^{-03}	TP53INP1
chr19:47852470-47852595	3	1.36 x 10^{-04} (2.33 x 10^{-05})	5.86	4.637 x 10^{-09}	1.873 x 10^{-03}	DHX34
chr6:30614168-30614422	7	2.51 x 10^{-04} (4.29 x 10^{-05})	5.85	5.019 x 10^{-09}	2.027 x 10^{-03}	C6orf136;
chr1:147801103-147801721	4	2.16 x 10^{-03} (3.79 x 10^{-04})	5.69	1.294 x 10^{-08}	5.227 x 10^{-03}	-
chr7:63505768-63506148	3	-1.77 x 10^{-03} (3.12 x 10^{-04})	-5.67	1.393 x 10^{-08}	5.628 x 10^{-03}	ZNF727
chr7:158110152-158110685	4	1.21 x 10^{-03} (2.14 x 10^{-04})	5.66	1.551 x 10^{-08}	6.264 x 10^{-03}	PTPRN2;
chr11:31391088-31391293	3	1.89 x 10^{-04} (3.39 x 10^{-05})	5.58	2.378 x 10^{-08}	9.606 x 10^{-03}	DCDC1; DNAJC24
chr13:47472050-47472140	4	-5.58 x 10^{-04} (1.04 x 10^{-04})	-5.39	6.932 x 10^{-08}	2.800 x 10^{-02}	HTR2A
chr7:73894884-73895109	3	-1.79 x 10^{-03} (3.33 x 10^{-04})	-5.39	7.224 x 10^{-08}	2.918 x 10^{-02}	GTF2IRD1
chr10:23982350-23982387	2	3.85 x 10^{-04} (7.22 x 10^{-05})	5.33	1.009 x 10^{-07}	4.075 x 10^{-02}	KIAA1217
chr11:63974153-63974229	3	1.55 x 10^{-04} (2.91 x 10^{-05})	5.32	1.018 x 10^{-07}	4.110 x 10^{-02}	FERMT3
chr1:20209848-20210269	2	-1.09 x 10^{-03} (2.05 x 10^{-04})	-5.32	1.055 x 10^{-07}	4.261 x 10^{-02}	OTUD3
chr7:54609587-54609671	2	2.51 x 10^{-04} (4.72 x 10^{-05})	5.31	1.112 x 10^{-07}	4.490 x 10^{-02}	VSTM2A

Appendix S

S1 Top 5 GO terms and KEGG pathways for CpGs in DMRs (BP = biological process; MF = molecular function; CC = cell compartment)

Model	Ontology/ KEGG	Term/Pathway	ID	N CpGs differentially methylated	N CpGs in term/pathway	P-value for enrichment	FDR corrected P-value
Total caffeine							
	BP	cell-cell adhesion via plasma-membrane adhesion molecules	GO:0098742	5	260	0.0001	1
	BP	homophilic cell adhesion via plasma membrane adhesion molecules	GO:0007156	4	160	0.0002	1
	MF	calcium ion binding	GO:0005509	6	670	0.0006	1
	CC	integral component of plasma membrane	GO:0005887	8	1520	0.001	1
	CC	intrinsic component of plasma membrane	GO:0031226	8	1595	0.002	1
	KEGG	Glycosphingolipid biosynthesis - ganglio series	path:hsa00604	1	15	0.033	1
	KEGG	Neuroactive ligand-receptor interaction	path:hsa04080	2	317	0.048	1
	KEGG	Various types of N-glycan biosynthesis	path:hsa00513	1	37	0.052	1
	KEGG	Circadian rhythm	path:hsa04710	1	30	0.055	1
	KEGG	Hippo signaling pathway - multiple species	path:hsa04392	1	28	0.073	1
Any vs. no Caffeine							
	MF	sarcosine dehydrogenase activity	GO:0008480	1	1	0.001	1
	BP	sarcosine catabolic process	GO:1901053	1	1	0.001	1
	MF	rRNA primary transcript binding	GO:0042134	1	1	0.001	1
	MF	oxidoreductase activity, acting on the CH-NH group of donors, flavin as acceptor	GO:0046997	1	2	0.001	1
	CC	endoplasmic reticulum	GO:0005783	5	1287	0.0010	1

	KEGG	Glycine, serine and threonine metabolism	path:hsa00260	1	35	0.019	1
	KEGG	Fatty acid degradation	path:hsa00071	1	41	0.024	1
	KEGG	Fatty acid metabolism	path:hsa01212	1	55	0.041	1
	KEGG	PPAR signaling pathway	path:hsa03320	1	72	0.045	1
	KEGG	Adipocytokine signaling pathway	path:hsa04920	1	66	0.051	1
Coffee							
	BP	mesenchymal stem cell migration	GO:1905319	1	3	0.002	1
	BP	regulation of mesenchymal stem cell migration	GO:1905320	1	3	0.002	1
	BP	positive regulation of mesenchymal stem cell migration	GO:1905322	1	3	0.002	1
	MF	C-X-C chemokine binding	GO:0019958	1	6	0.003	1
	MF	N-acetyl-beta-glucosaminyl-glycoprotein 4-beta-N-acetylgalactosaminyltransferase activity	GO:0033842	1	2	0.003	1
	KEGG	Autophagy - other	path:hsa04136	1	29	0.019	1
	KEGG	Various types of N-glycan biosynthesis	path:hsa00513	1	37	0.023	1
	KEGG	Viral protein interaction with cytokine and cytokine receptor	path:hsa04061	1	95	0.031	1
	KEGG	Th1 and Th2 cell differentiation	path:hsa04658	1	87	0.063	1
	KEGG	Autophagy - animal	path:hsa04140	1	129	0.081	1
Tea							
	MF	nicotinate phosphoribosyltransferase activity	GO:0004516	1	1	0.0004	1
	BP	nicotinate nucleotide biosynthetic process	GO:0019357	1	1	0.0004	1
	BP	nicotinate nucleotide salvage	GO:0019358	1	1	0.0004	1
	BP	pyridine nucleotide salvage	GO:0019365	1	1	0.0004	1
	BP	NAD salvage	GO:0034355	1	1	0.0004	1
	KEGG	Nicotinate and nicotinamide metabolism	path:hsa00760	1	34	0.022	1
	KEGG	Taste transduction	path:hsa04742	1	82	0.066	1
	KEGG	GnRH secretion	path:hsa04929	1	62	0.098	1
	KEGG	ECM-receptor interaction	path:hsa04512	1	86	0.119	1
	KEGG	GABAergic synapse	path:hsa04727	1	84	0.125	1
Cola							
	MF	alpha-1,6-mannosylglycoprotein 2-beta-N-acetylglucosaminyltransferase activity	GO:0008455	1	1	0.0003	1

	BP	thyroid-stimulating hormone secretion	GO:0070460	1	1	0.002	1
	BP	regulation of thyroid-stimulating hormone secretion	GO:2000612	1	1	0.002	1
	CC	late recombination nodule	GO:0005715	1	1	0.002	1
	BP	meiotic metaphase I plate congression	GO:0043060	1	1	0.002	1
	MF	alpha-1,6-mannosylglycoprotein 2-beta-N-acetylglucosaminyltransferase activity	GO:0008455	1	1	0.0003	1
	KEGG	Mismatch repair	path:hsa03430	1	22	0.018	1
	KEGG	Aminoacyl-tRNA biosynthesis	path:hsa00970	1	43	0.029	1
	KEGG	Various types of N-glycan biosynthesis	path:hsa00513	1	37	0.032	1
	KEGG	N-Glycan biosynthesis	path:hsa00510	1	48	0.036	1
	KEGG	Fanconi anemia pathway	path:hsa03460	1	48	0.037	1
Total caffeine – female sex							
	CC	nucleosome	GO:0000786	2	90	0.0017	1
	MF	endoribonuclease inhibitor activity	GO:0060698	1	2	0.002	1
	BP	negative regulation of ribonuclease activity	GO:0060701	1	2	0.002	1
	BP	negative regulation of endoribonuclease activity	GO:0060702	1	2	0.002	1
	CC	DNA packaging complex	GO:0044815	2	98	0.002	1
	KEGG	Systemic lupus erythematosus	path:hsa05322	2	113	0.003	0.93
	KEGG	Ubiquitin mediated proteolysis	path:hsa04120	2	132	0.006	0.93
	KEGG	Alcoholism	path:hsa05034	2	164	0.008	0.93
	KEGG	Various types of N-glycan biosynthesis	path:hsa00513	1	37	0.033	1
	KEGG	Progesterone-mediated oocyte maturation	path:hsa04914	1	86	0.078	1
Total caffeine – male sex							
	BP	positive regulation of phosphatidylinositol biosynthetic process	GO:0010513	2	5	1.2×10^{-05}	0.275
	BP	regulation of phosphatidylinositol biosynthetic process	GO:0010511	2	7	4.2×10^{-05}	0.480
	BP	positive regulation of phospholipid biosynthetic process	GO:0071073	2	11	7.2×10^{-05}	0.544
	BP	regulation of phospholipid biosynthetic process	GO:0071071	2	18	$2.4\ 10^{-04}$	1.000
	BP	response to methyl methanesulfonate	GO:0072702	1	1	6.9×10^{-04}	1.000

Appendix T

T1 ALSPAC definition of variables and covariates

- **Offspring internalising problems:** In ALSPAC, internalising problems at the age 3 and 7 were assessed through questionnaires at 47 months and 81 months, respectively. The total emotional symptoms score was calculated by summing the score of each item for children that had complete data on all 5 items. To obtain the item mean, the total score was divided by 5 (the number of items of the emotional symptoms score)
- **Offspring age:** Offspring's age at completion of the questionnaires was first generated based on offspring's date of birth and the date of completion of the questionnaires. Offspring's age was transformed from days to weeks by dividing age in days by 7.
- **Maternal education** (as proxy for maternal socioeconomic position): Maternal education was assessed in week 32 of gestation and coded as an ordinal variable: "Vocational/CSE" = 1, "O level" (at 16, equivalent to lower grades of ordinary-level) = 2, "A level" (ordinary-level school-leaving certificate (at 16) = 3, and "Degree" (advanced-level school-leaving certificate (post-16)/degree) = 4.
- **Maternal smoking during pregnancy** was assessed as an ordinary variable representing 0 = no or early smoking during pregnancy, 1 = Stopped before the second trimester of pregnancy and 2 = Smoking in the third trimester or throughout pregnancy.
- **Maternal age** continuous numeric variable in years assessed at birth of study child.
- **Parity** has been assessed at 18 weeks gestation as number of previous pregnancies resulting in either a livebirth or a stillbirth.
- **Gestational age** was calculated (in days) based on the date of the mother's last menstrual period (LMP) when the mother was certain of this, but for uncertain LMPs and conflicts with clinical assessment the ultrasound assessment was used. Where maternal report and ultrasound assessment conflicted, an experienced obstetrician reviewed clinical records and made a best estimate.
- **Maternal anxiety and depression during pregnancy:** Maternal anxiety and depressive symptoms during pregnancy were assessed with the Edinburgh Postnatal Depression scale in a questionnaire at 18 weeks gestation. A total score was generated if mothers had data on each of the ten items by summing up the individual item scores.
- **Offspring sex** was taken from obstetric records

T2 Generation R definition of variables and covariates

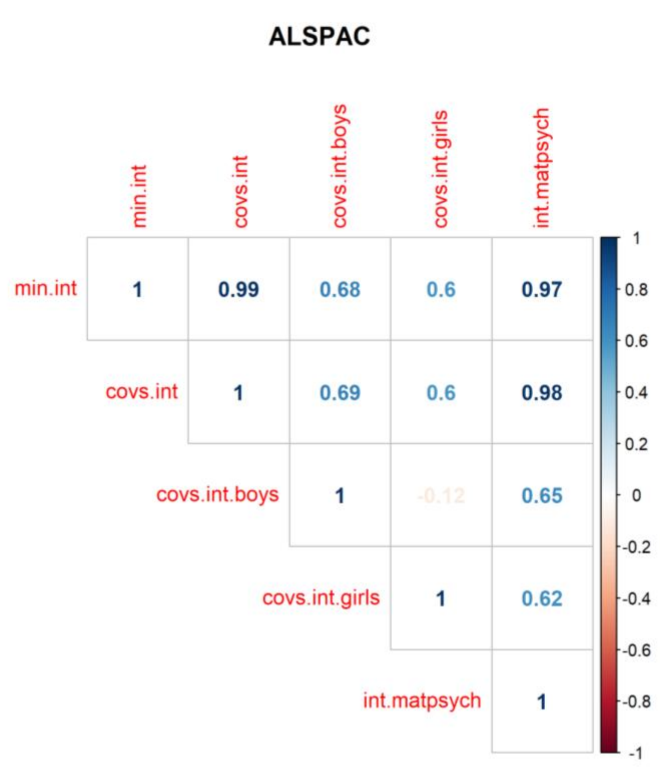
- **Offspring internalising problems:** Internalising problems at the age 3 and 7 were assessed through questionnaires at 36 months and 72 months, respectively. The CBCL internalising scale score was generated according to the CBCL manual.
- **Offspring age** at internalising problems assessment was reported by mothers in the corresponding questionnaire.
- **Maternal education** (as proxy for maternal socioeconomic position): Maternal education was assessed in week 12-20 of gestation and coded as an ordinal variable: " **“**no education finished " = 1, “primary” (at 12, equivalent to lower grades of ordinary-level) = 2, "secondary, phase 1" (ordinary-level school-leaving certificate (at 16) = 3, and “3-secondary, phase 2” (advanced-level school-leaving certificate (post-16)) = 4, “higher, phase 2” (education after High School) = 5.
- **Maternal smoking during pregnancy** was assessed as an ordinary variable representing 0 = no or early smoking during pregnancy, 1 = Stopped before the second trimester of pregnancy and 2 = Smoking in the third trimester or throughout pregnancy.
- **Maternal age** continuous numeric variable in years assessed at study intake.
- **Gestational age** was assessed at birth of study child.
- **Maternal anxiety and depression during pregnancy** was assessed through questionnaires at weeks 20-25 of gestation using the depression and anxiety subscales of the global severity index (GSI)**.** The GSI assesses severity of anxiety and depressive symptoms over the past 7 days. A combined score was generated through taking the mean of the depression and anxiety subscales.
- **Offspring sex** was derived from birth records.

T3 MoBa definition of variables and covariates

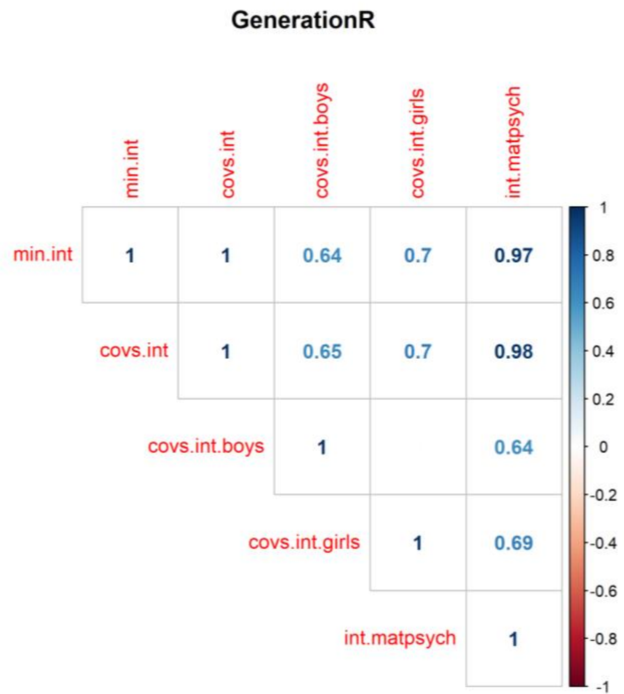
- **Offspring internalising problems:** Internalising problems at the age 3 were assessed through a questionnaire at 36 months. Nine selected items of the original CBCL internalising subscale were used to calculate a total internalising problems score by summing the score of each item for children that had complete data on all 9 items.
- **Offspring age:** Offspring's age at completion of the questionnaires was first generated based on offspring's date of birth and the date of completion of the questionnaires.
- **Maternal education** (as proxy for maternal socioeconomic position): Maternal education was assessed at 15 weeks gestation and coded as an ordinal variable: "<High Sch " = 0, "High School degree" = 1, " Some college " = 2, and "+4yr College" = 3.
- **Maternal smoking during pregnancy** was assessed as an ordinary variable representing 0 = no or early smoking during pregnancy, 1 = Stopped before the second trimester of pregnancy and 2 = Smoking in the third trimester or throughout pregnancy.
- **Maternal age** continuous numeric variable in years collected from Norwegian Medical Birth Registry.
- **Parity** has been assessed through linkage to the national health registries**.**
- **Gestational age** has been assessed through linkage to the national health registries
- **Maternal anxiety and depression during pregnancy:** Maternal anxiety and depressive symptoms during pregnancy were assessed with selective items from the (Hopkins) Symptoms Checklist-25 (SCL-25). The original SCL-25 scale comprises Maternal anxiety was assessed with 10 items for anxiety and 15 items that assess depression that are rated on a 4-point Likert scale (from "not at all, bothered," to "extremely bothered"). Mothers reported on their anxiety and depressive symptoms in a questionnaire at gestational week 15.
- **Offspring sex** has been assessed through linkage to the national health registries.

Appendix U

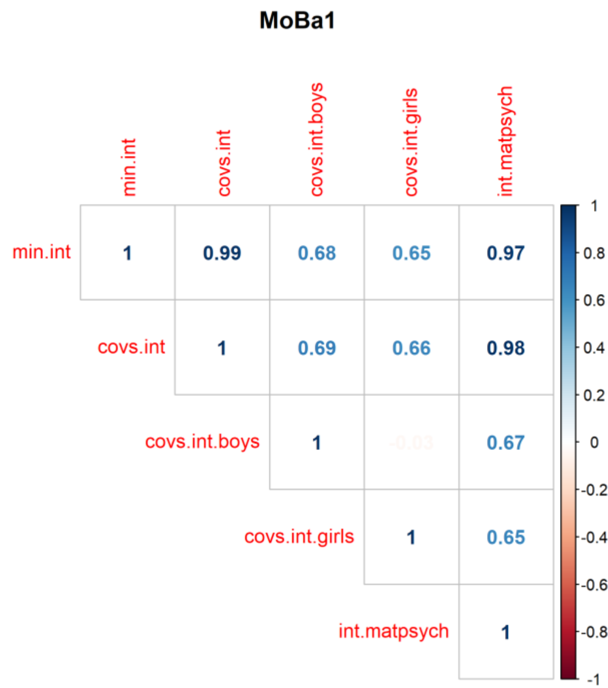
U1 Correlation plot cord blood meta-analysis age 3 – ALSPAC



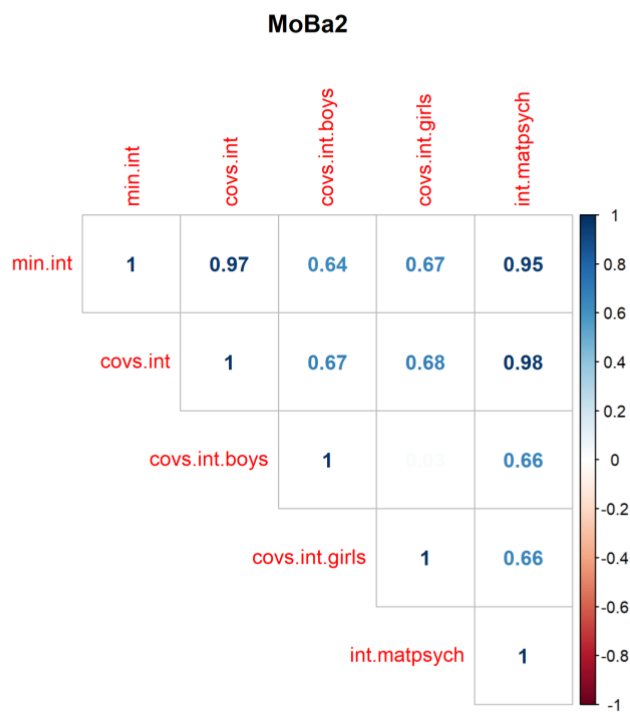
U2 Correlation plot cord blood meta-analysis age 3 – Generation



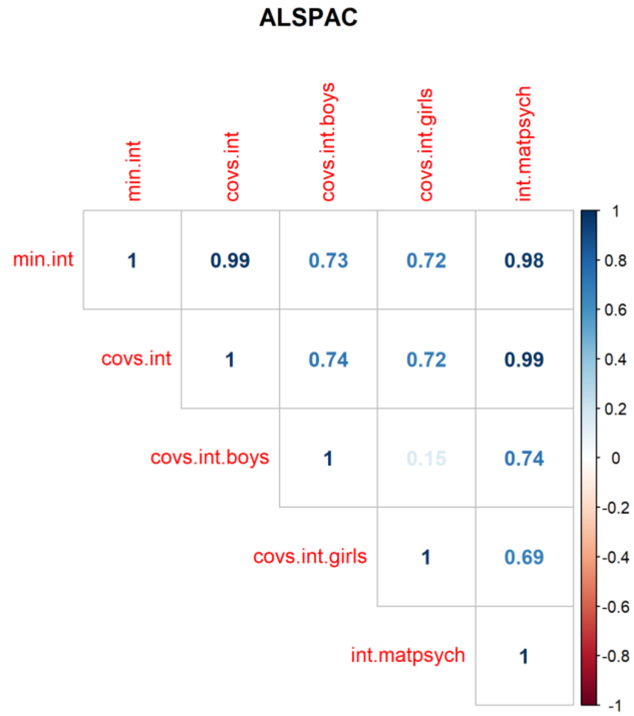
U3 Correlation plot cord blood meta-analysis age 3 – MoBa1



U4 Correlation plot cord blood meta-analysis age 3 – MoBa2

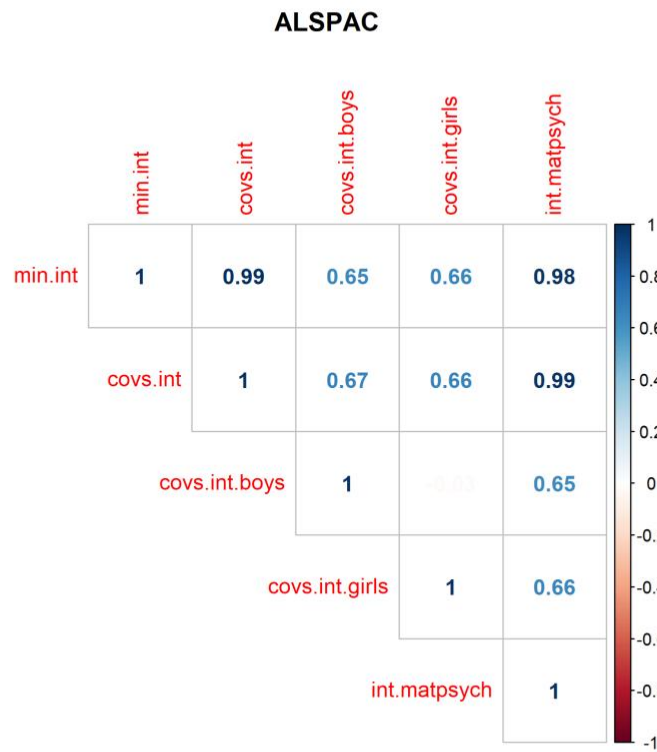


U5 Correlation plot cord blood meta-analysis age 7 – ALSPAC

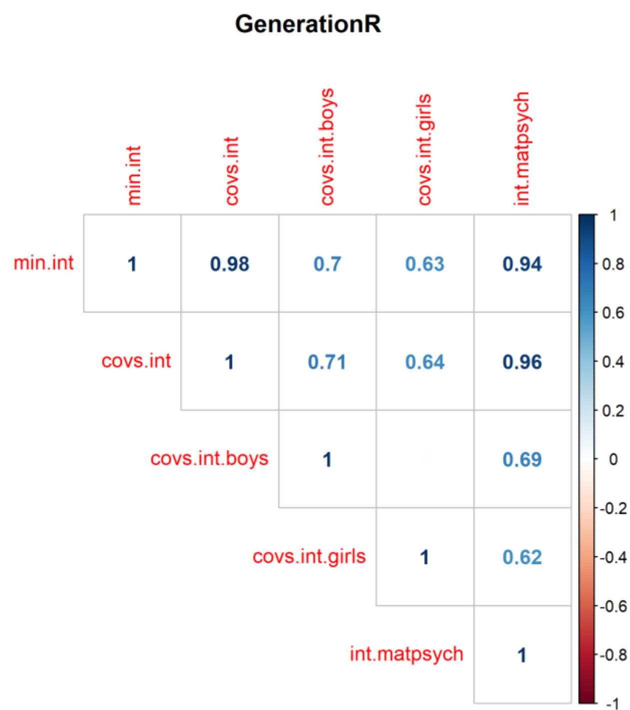


U6 Correlation plot cord blood meta-analysis age 7 – Generation R

U7 Correlation plot cross-sectional meta-analysis age 7 – ALSPAC

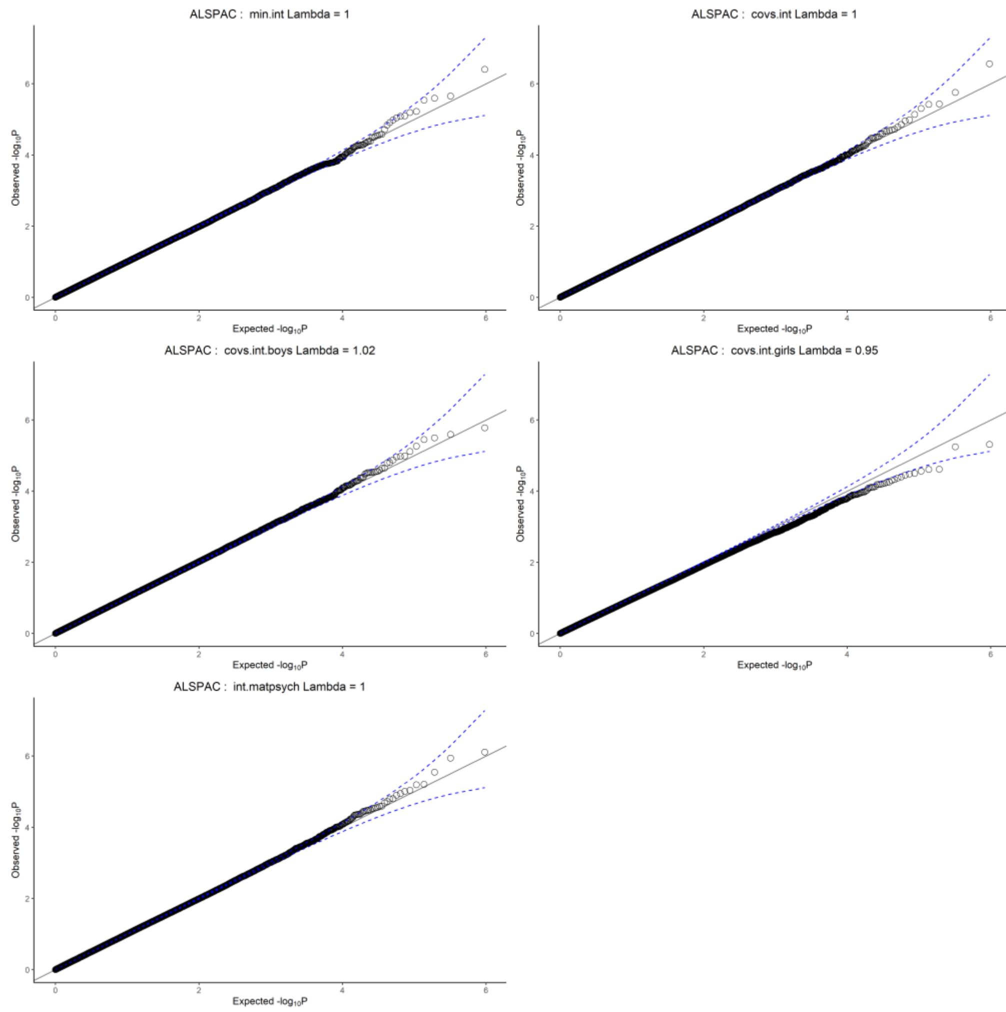


U8 Correlation plot cross-sectional meta-analysis age 7 – Generation R

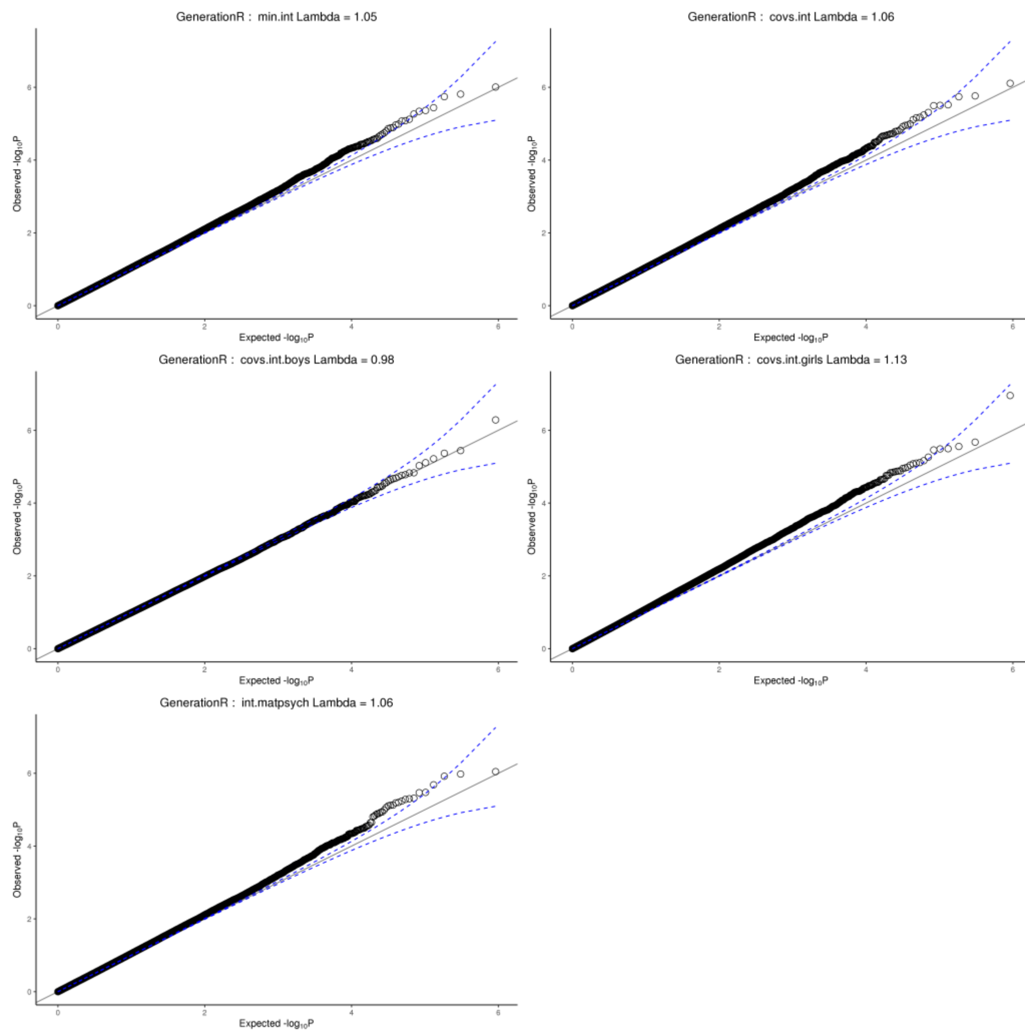


Appendix V

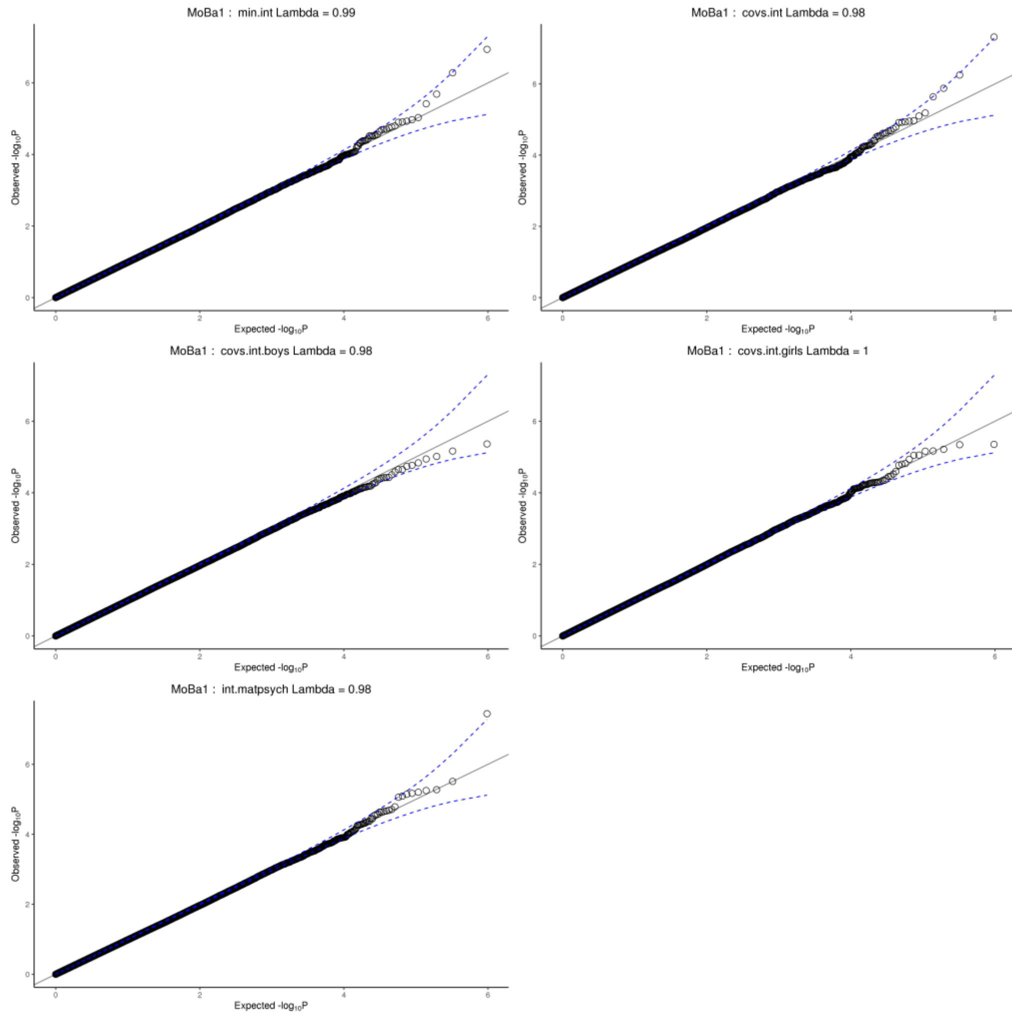
V1 QQ-plots cord blood meta-analysis age 3 – ALSPAC



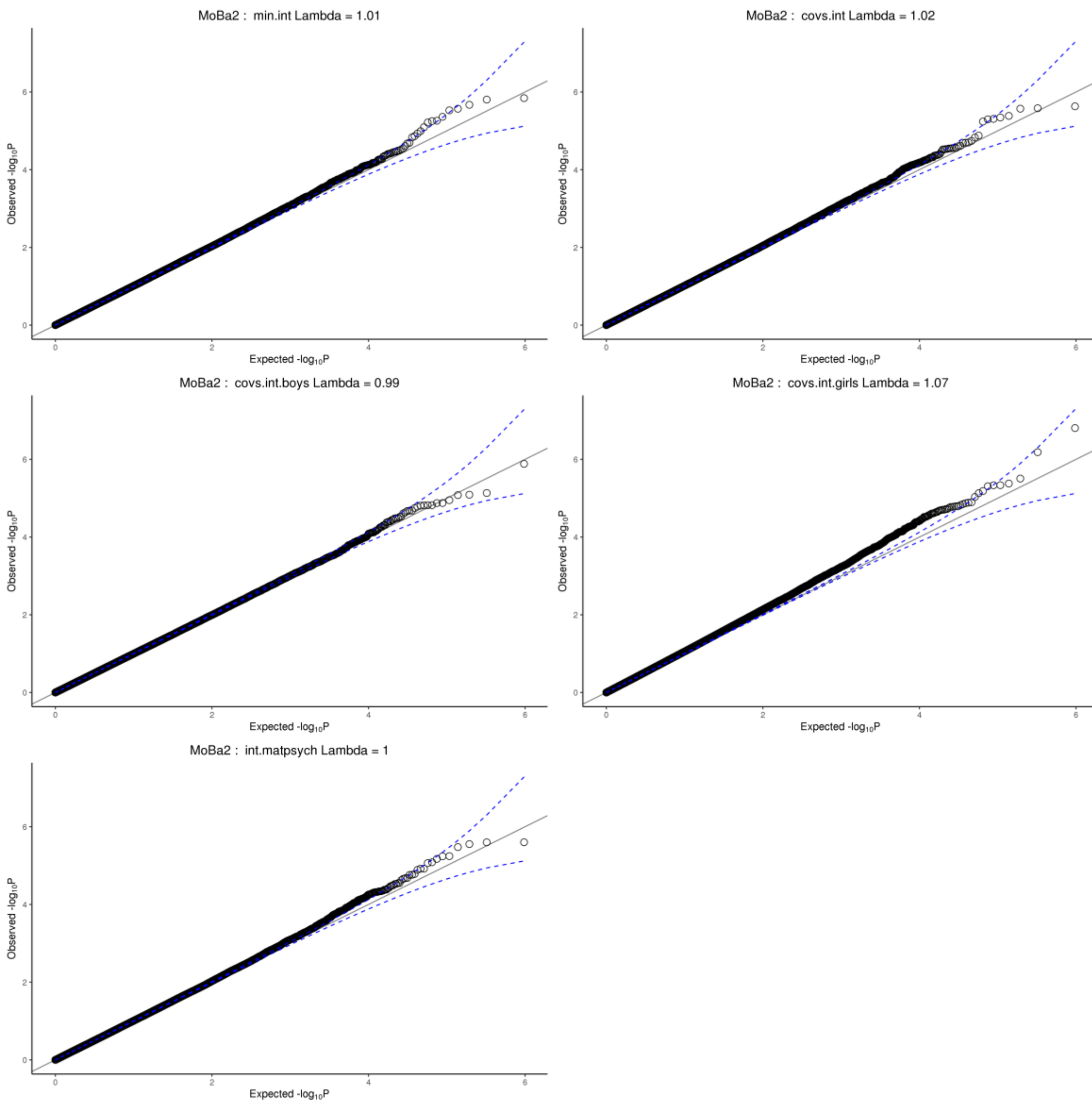
V2 QQ-plots cord blood meta-analysis age 3 – Generation R



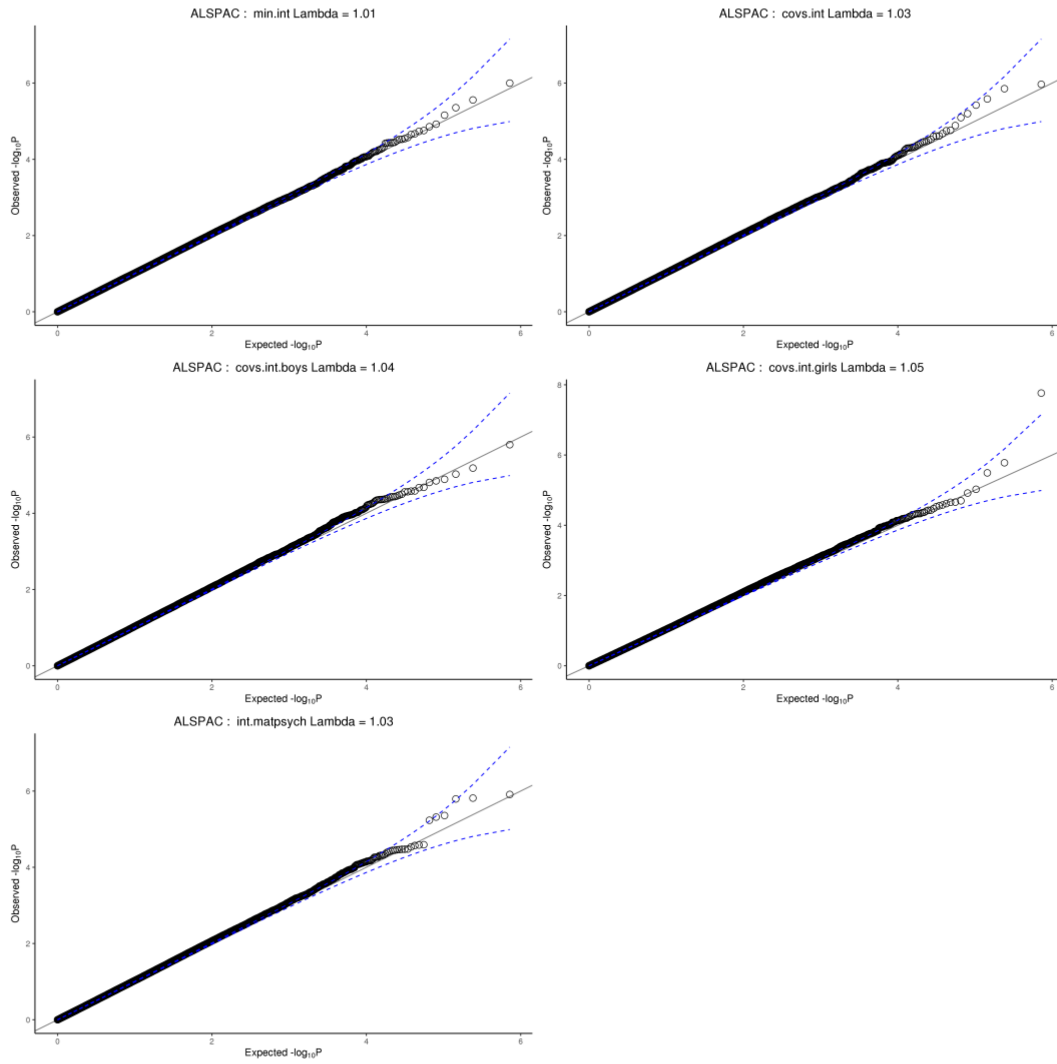
V3 QQ-plots cord blood meta-analysis age 3 – MoBa1



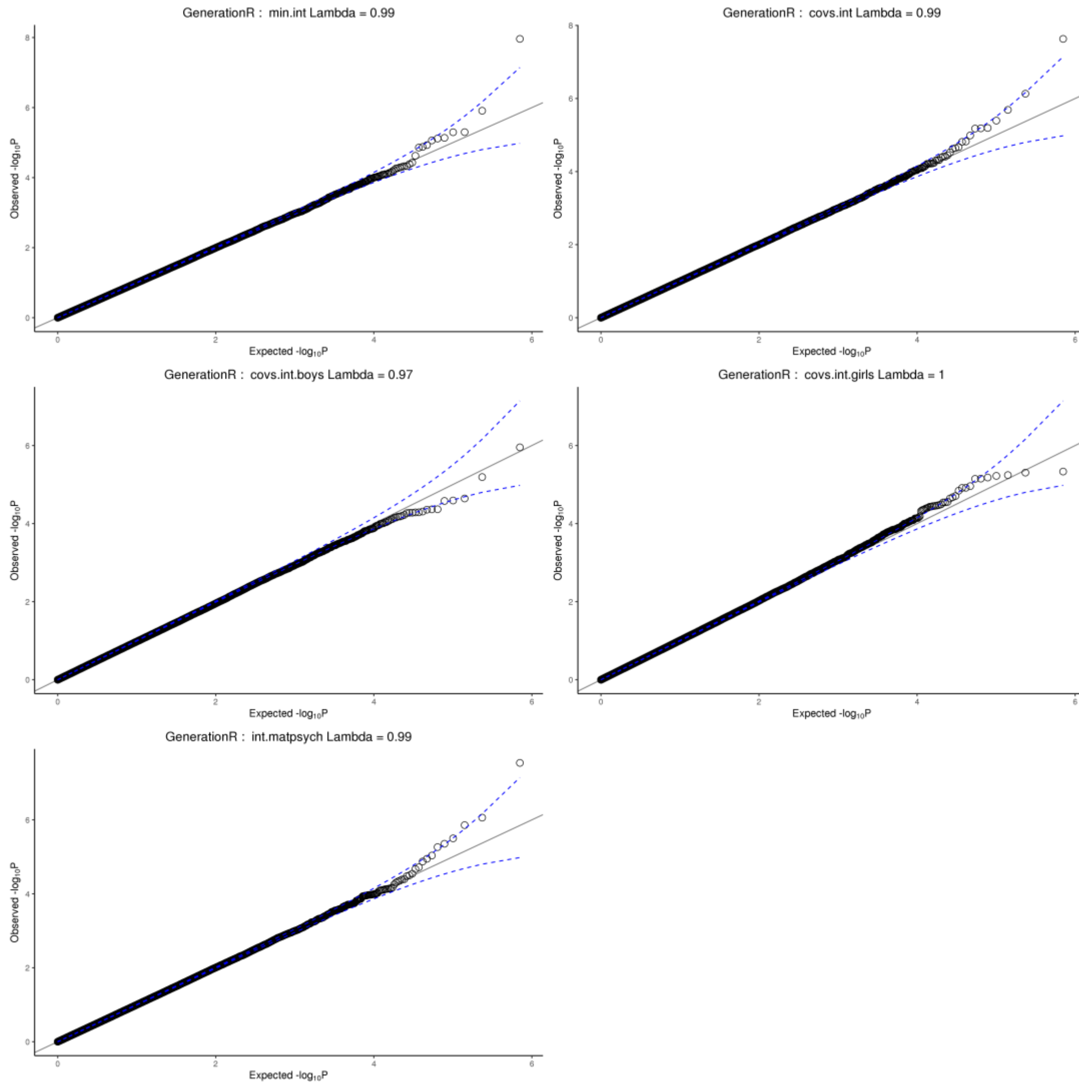
V4 QQ-plots cord blood meta-analysis age 3 MoBa2



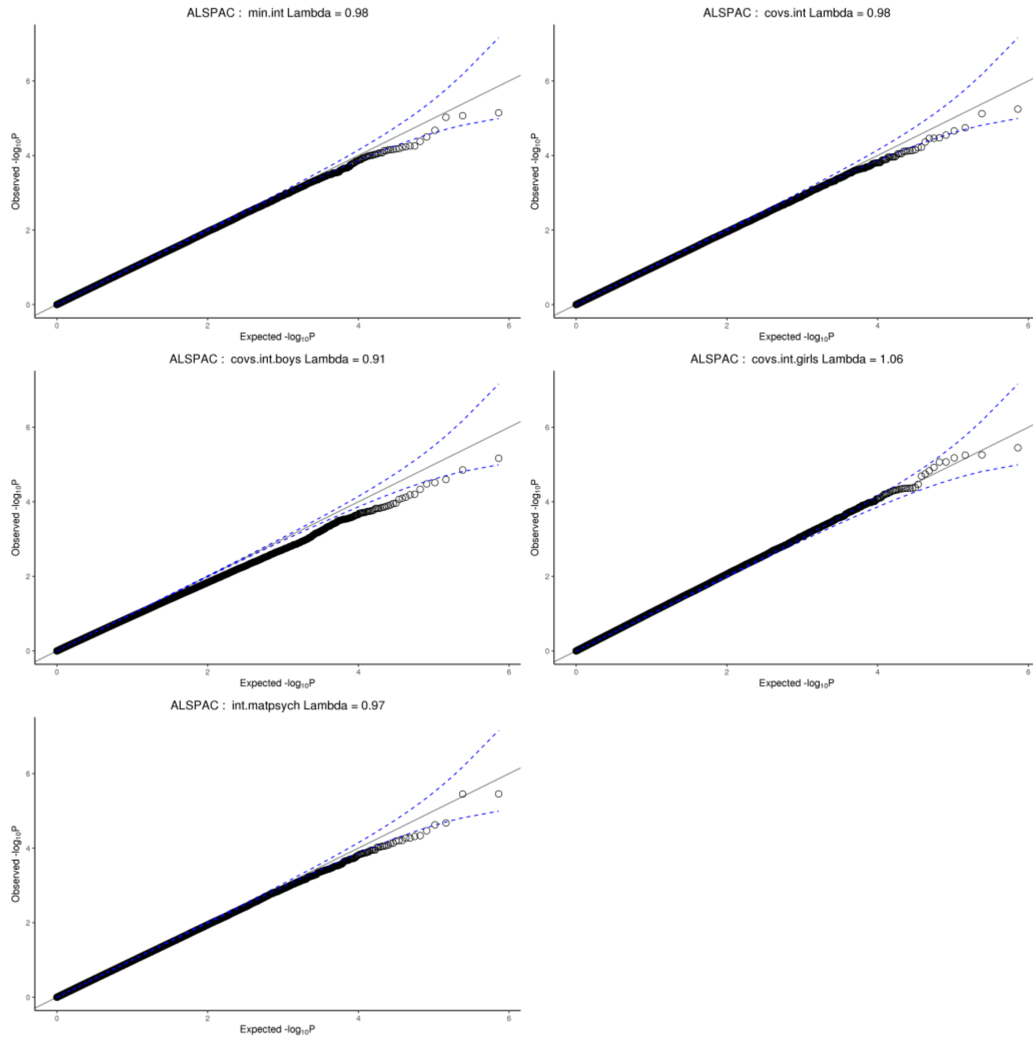
V5 QQ-plots of the cord blood meta-analysis age 7 – ALSPAC



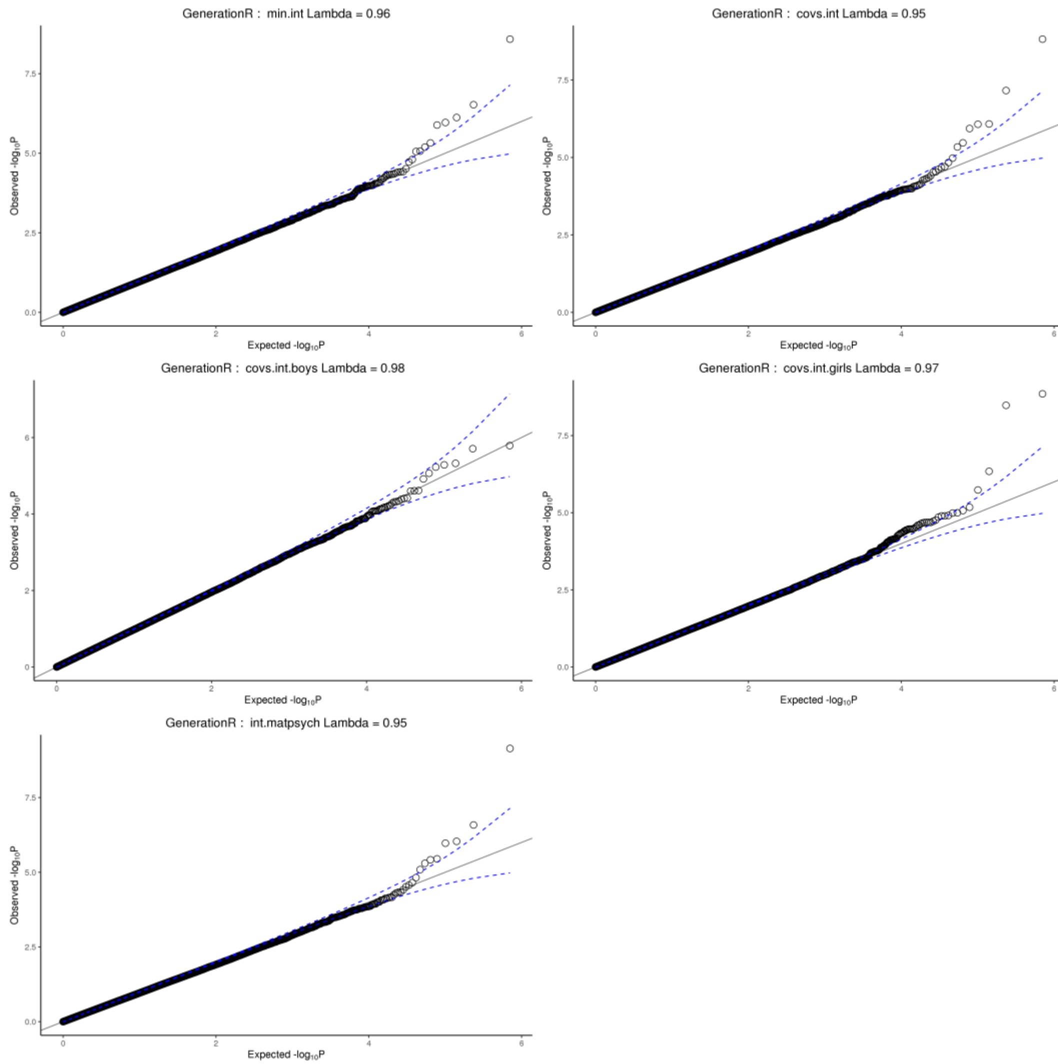
V6 QQ-plots of the cord blood meta-analysis age 7 – Generation R



V7 QQ-plots of the childhood peripheral blood meta-analysis age 7 – ALSPAC

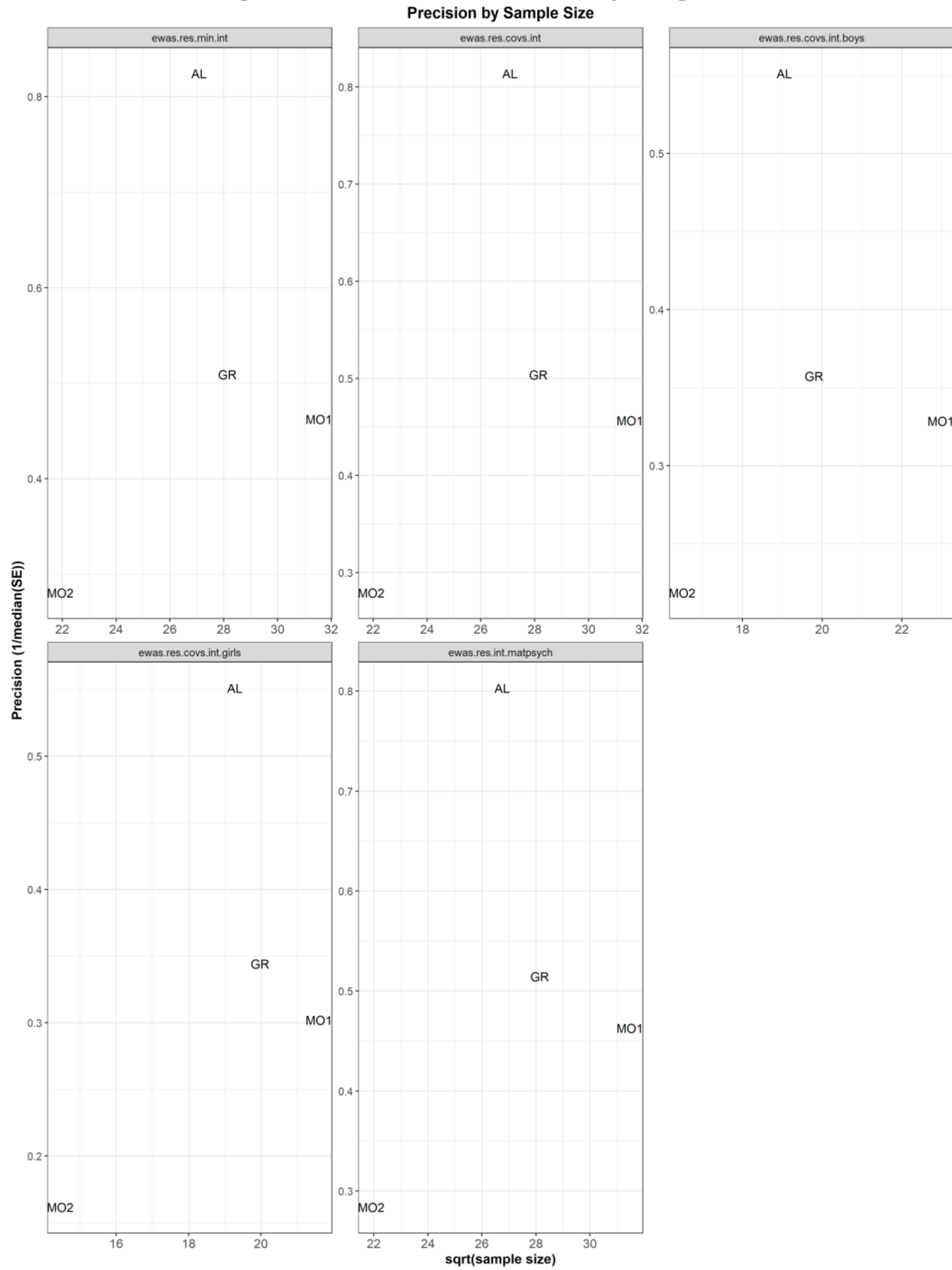


V8 QQ-plots of the childhood peripheral blood meta-analysis age 7 – Generation R

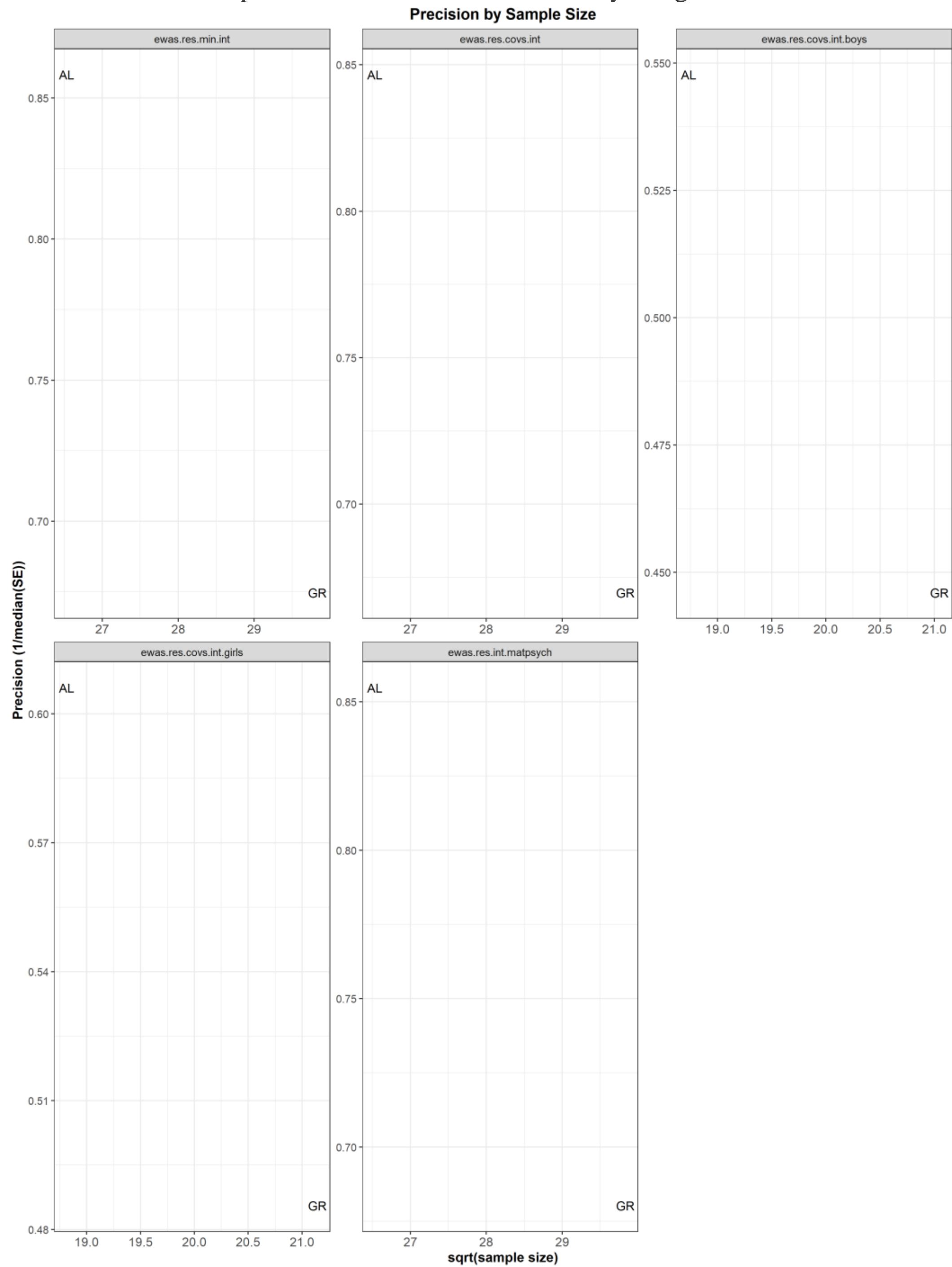


Appendix W

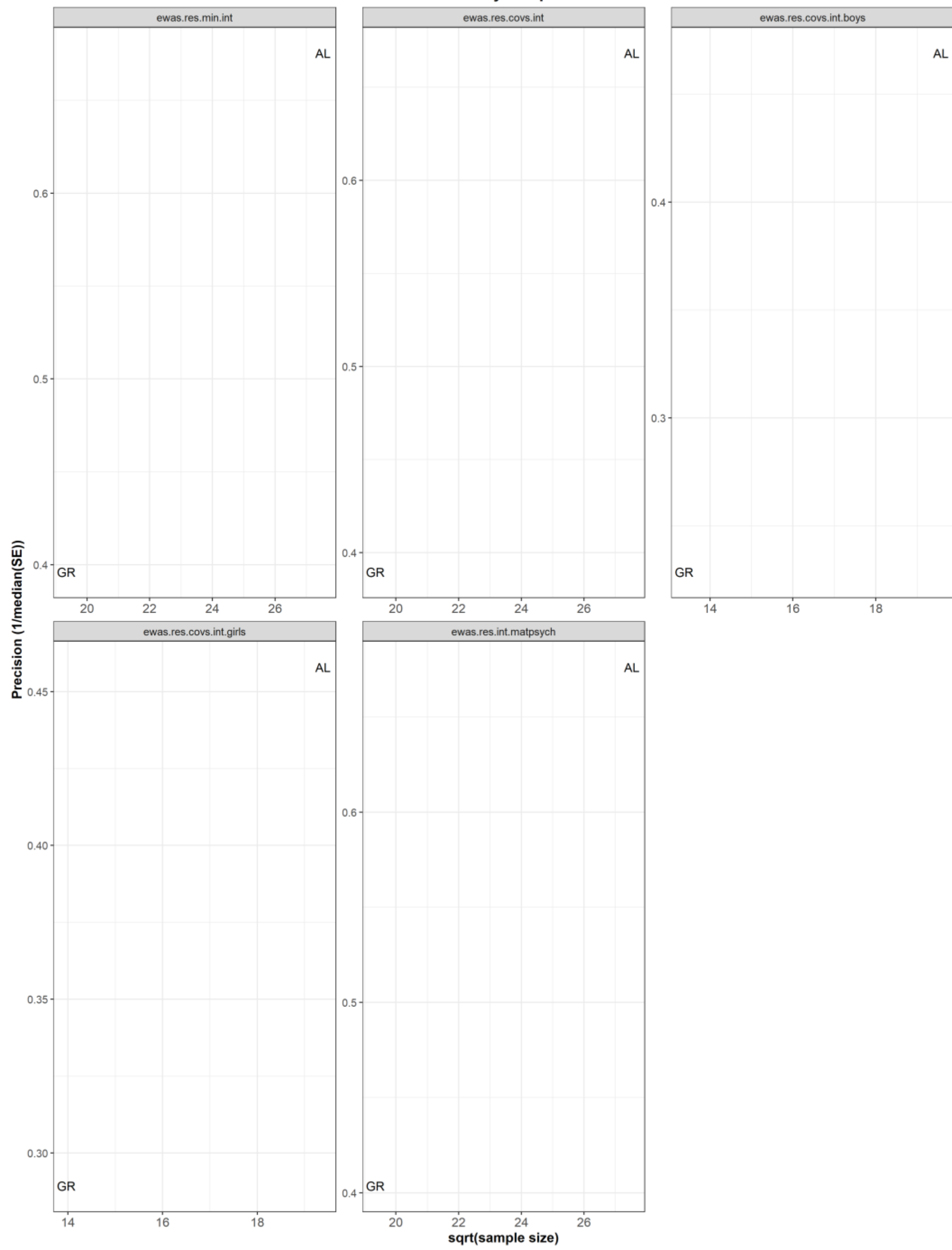
W1 Precision plot of the cord blood meta-analysis age 3



W2 Precision plot of the cord blood meta-analysis age 7

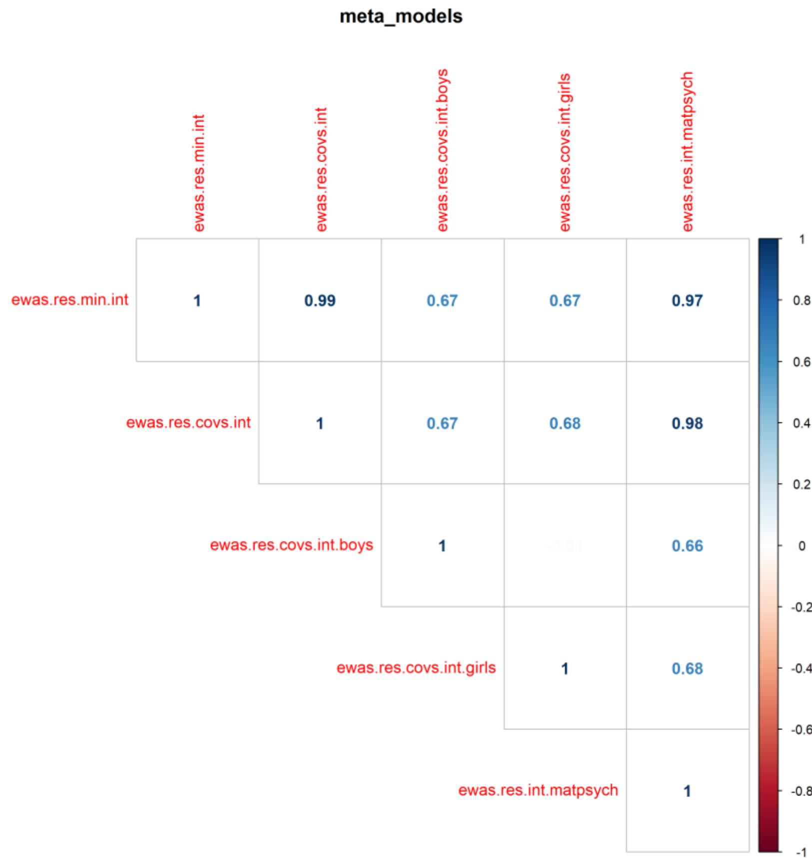


W3 Precision plot of the childhood peripheral blood meta-analysis age 7

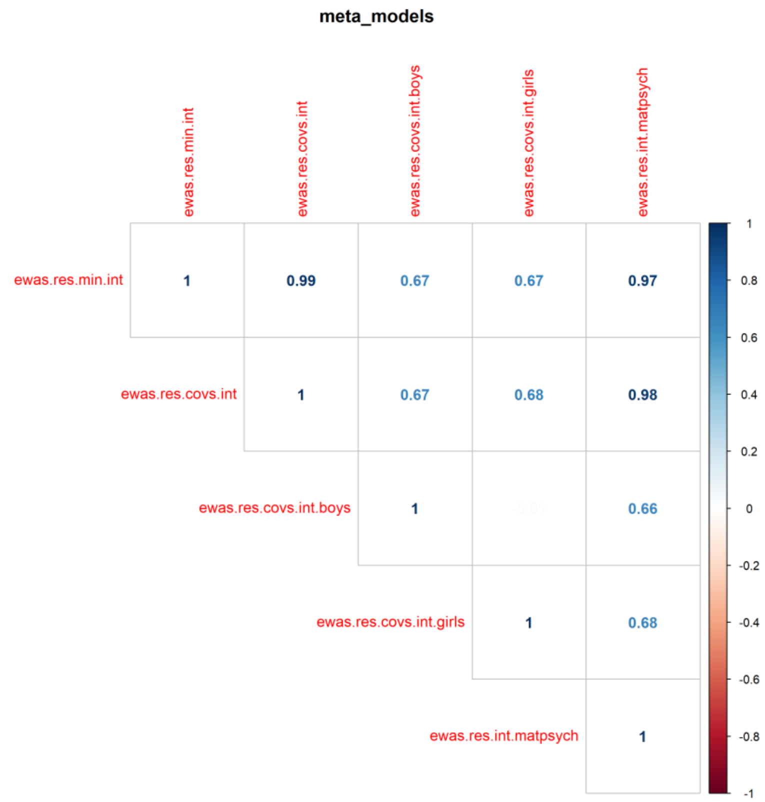


Appendix X

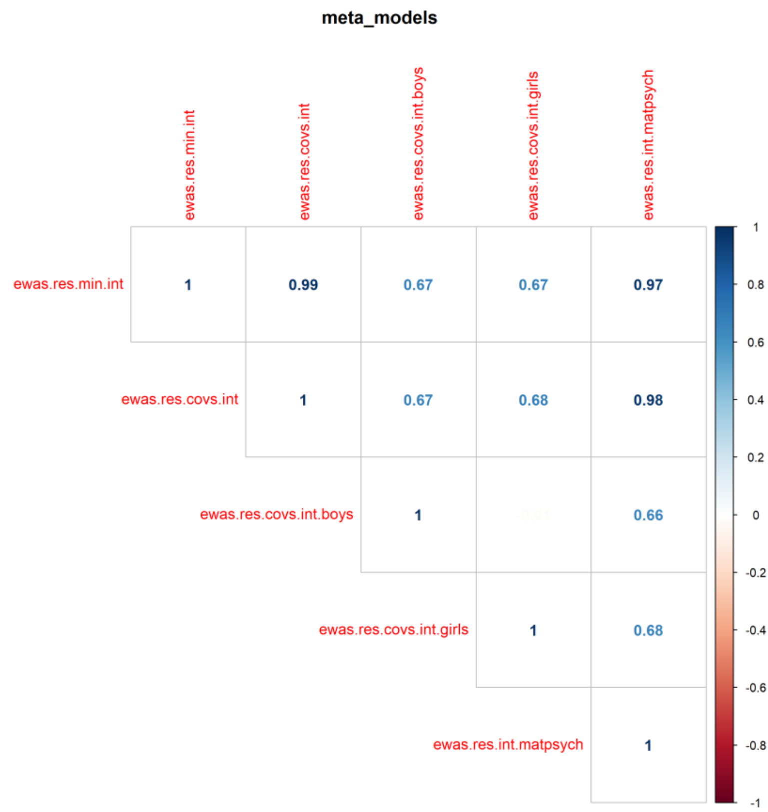
X1 Correlation matrix of the cord blood probe-level meta-analysis results: Internalising problems age 3



X2 Correlation matrix of the cord blood probe-level meta-analysis results: Internalising problems age 7

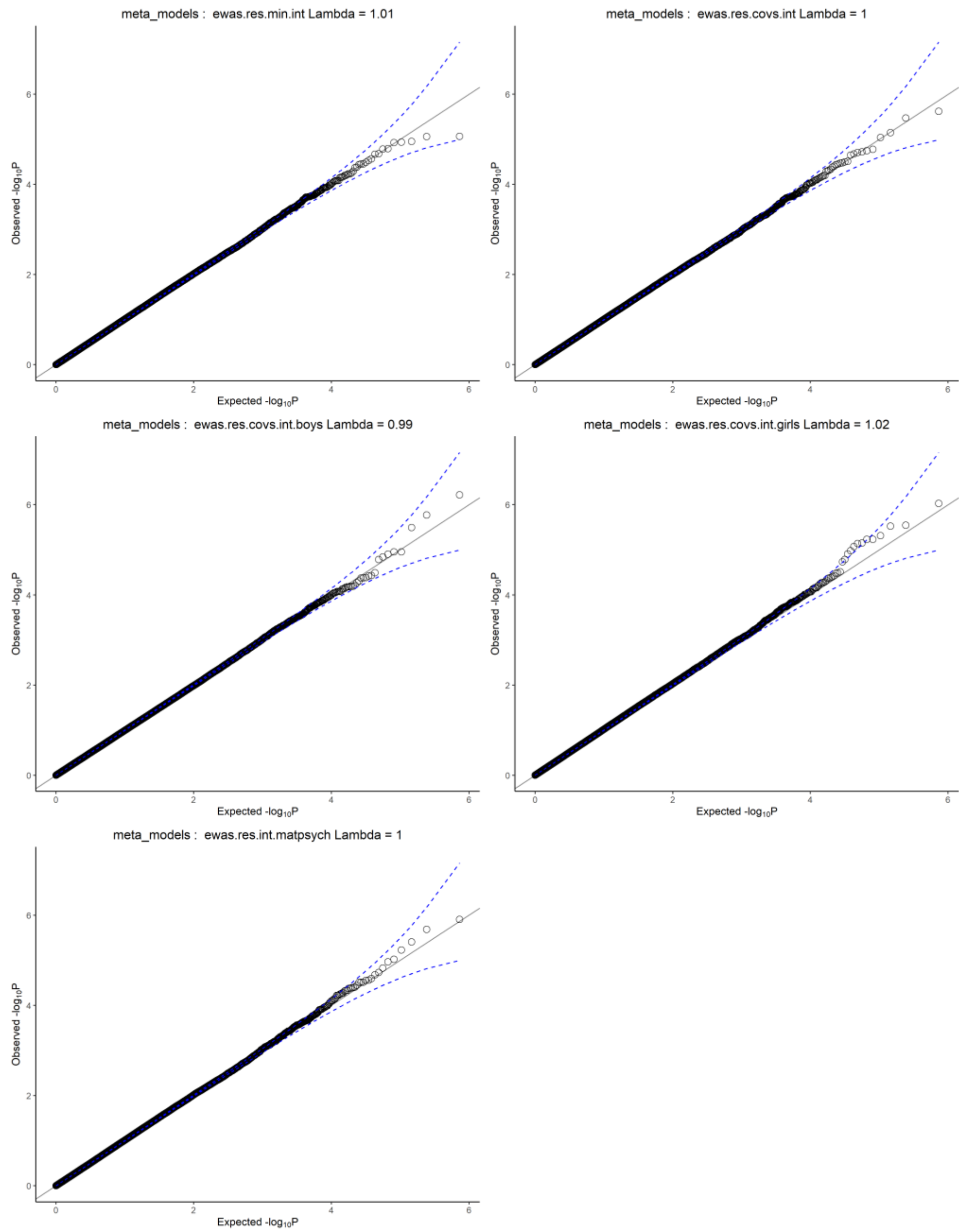


X3 Correlation matrix of the childhood peripheral blood probe-level meta-analysis results: Internalising problems age 7

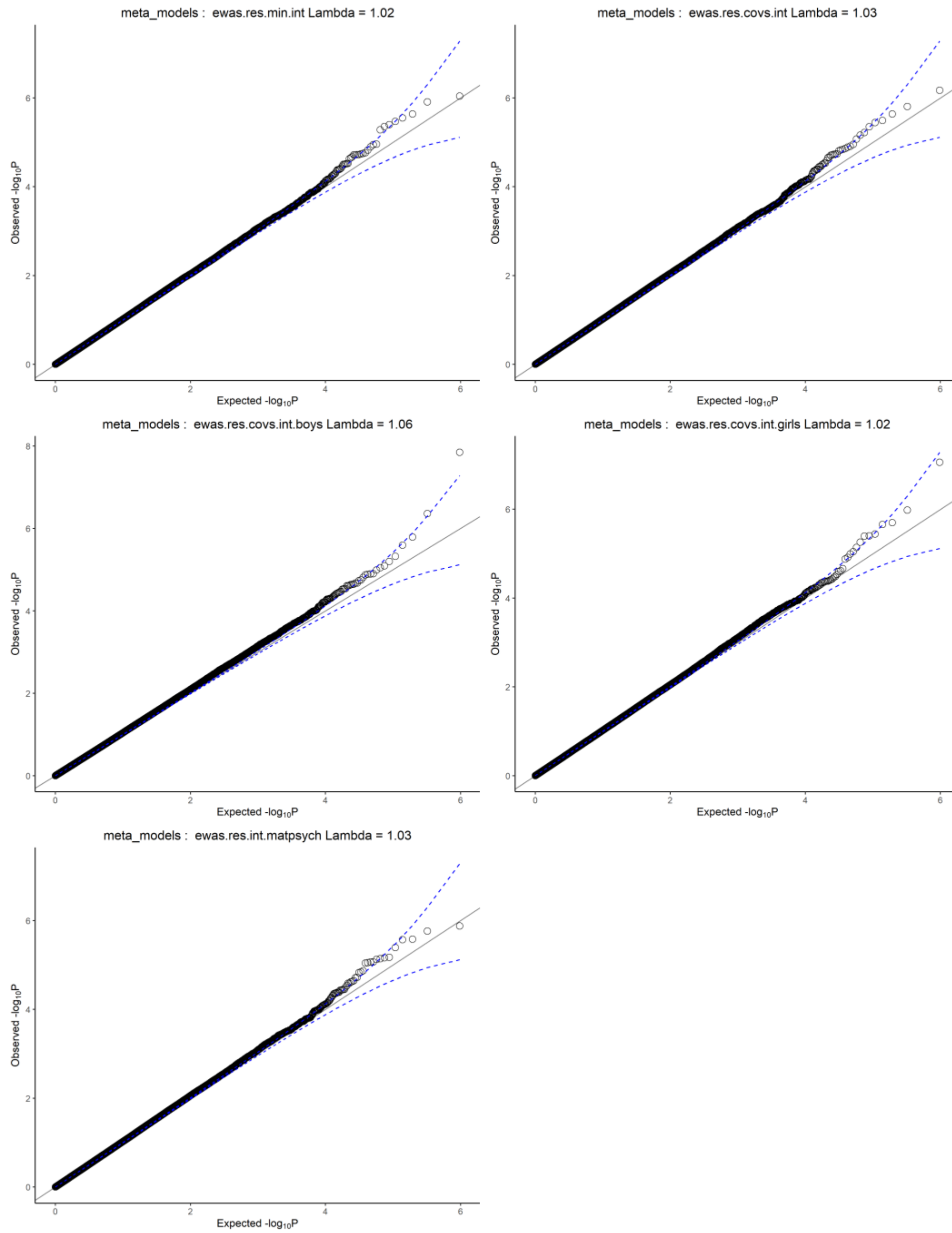


Appendix Y

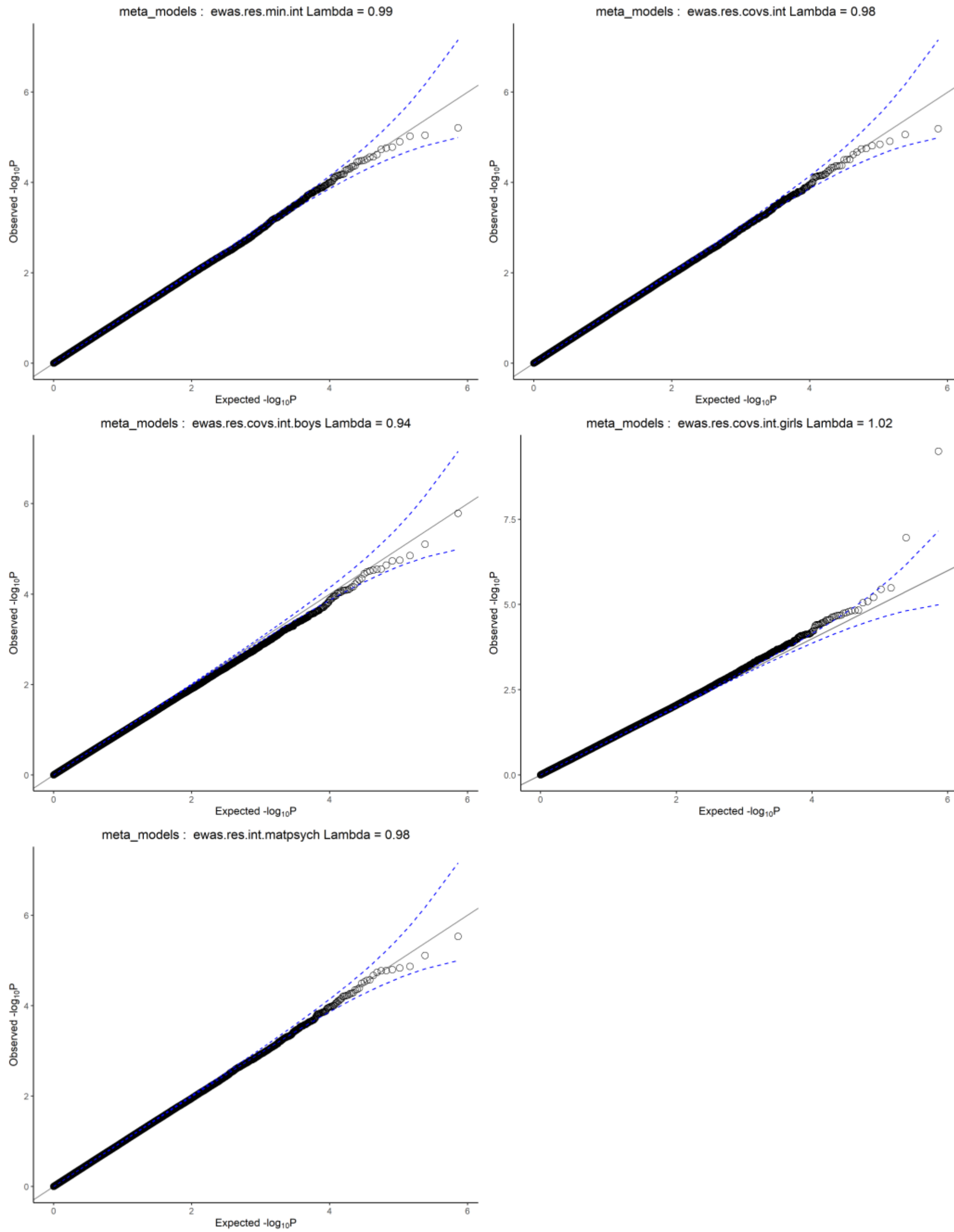
Y1 QQ-plot of the cord blood probe-level meta-analysis results: Internalising problems age 3



Y2 QQ-plot of the cord blood probe-level meta-analysis results: Internalising problems age 7

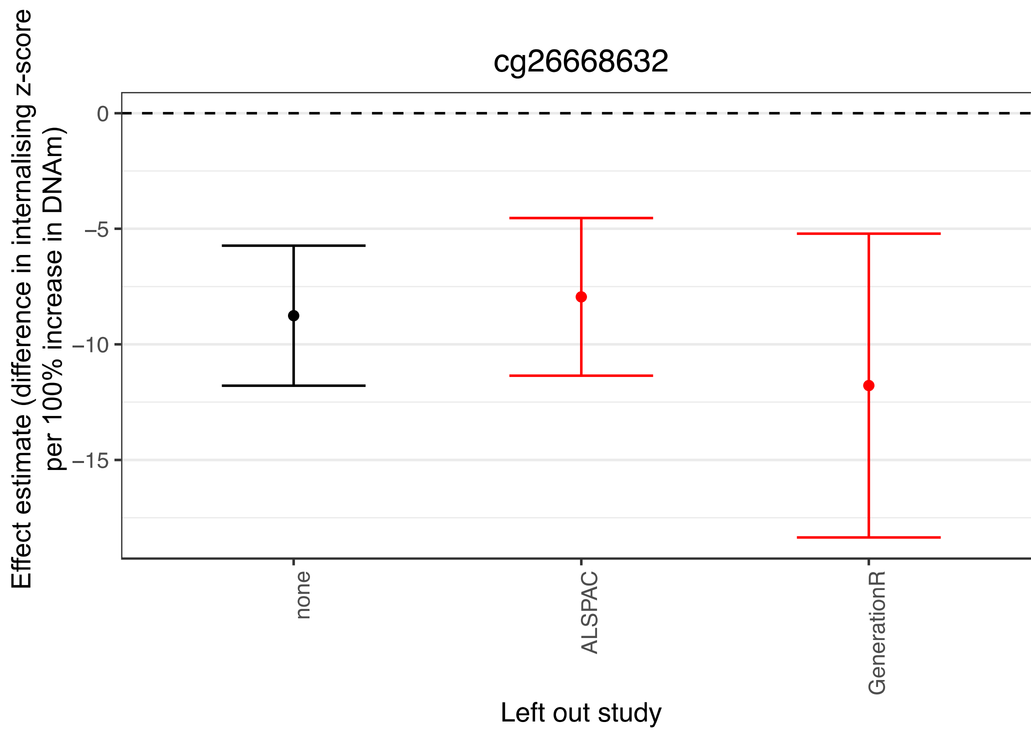


Y3 QQ-plot of the childhood peripheral blood probe-level meta-analysis results: Internalising problems age 7

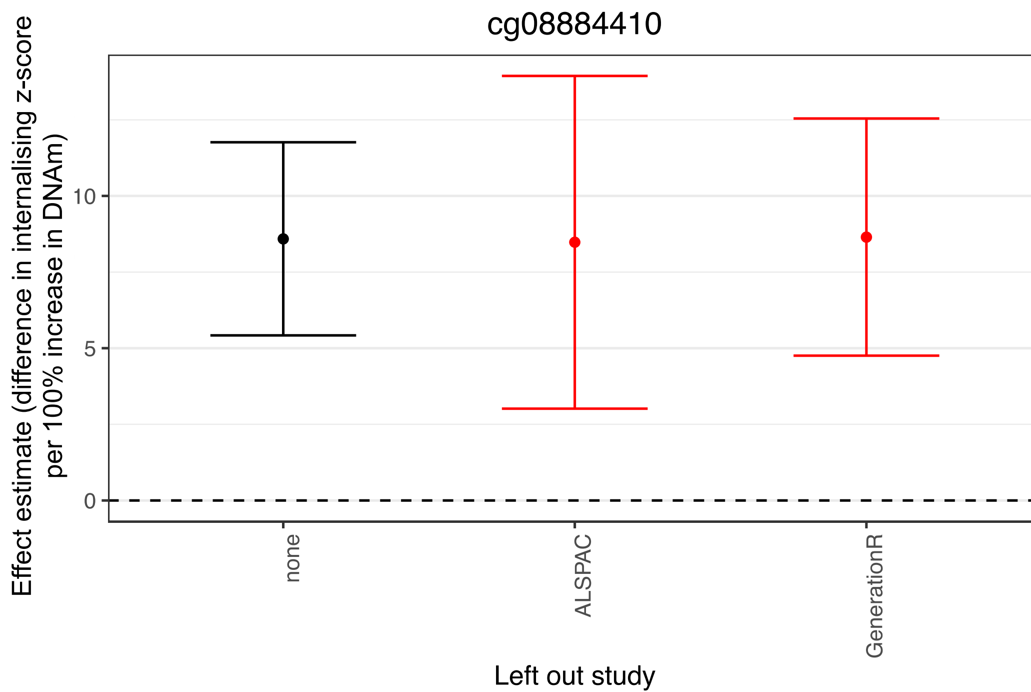


Appendix Z

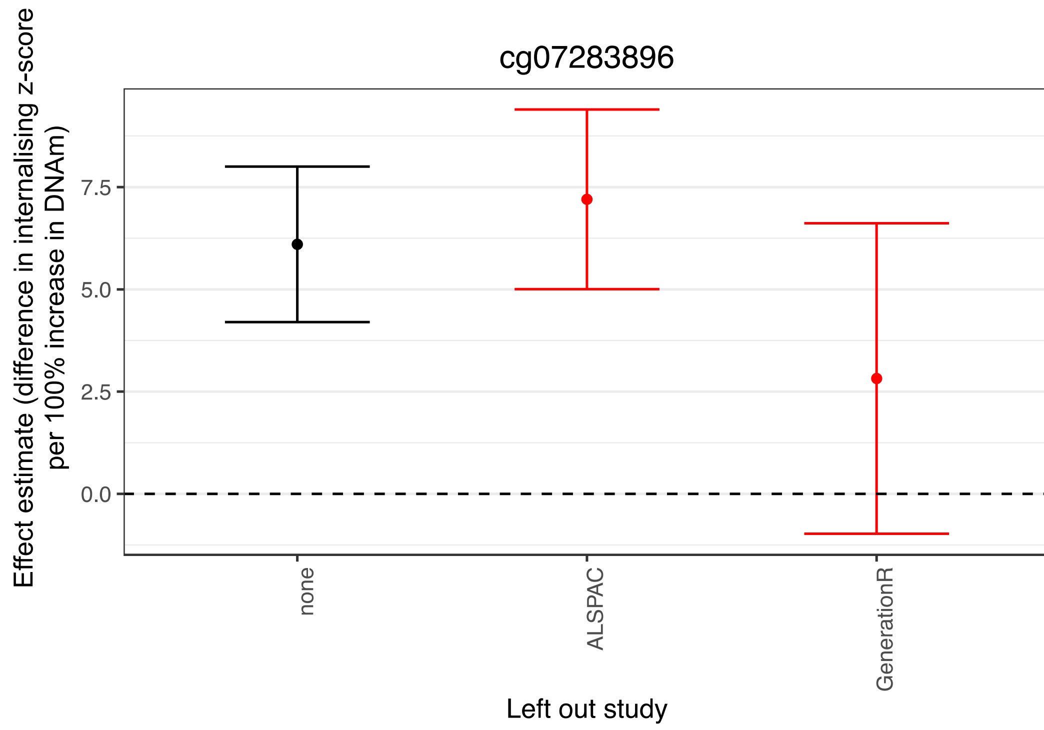
Z1 Leave-one-out plot of the cord-blood meta-analysis age 7 (Cg26668632)



Z2 Leave-one-out plot of the childhood peripheral blood meta-analysis age 7 (Cg08884410)

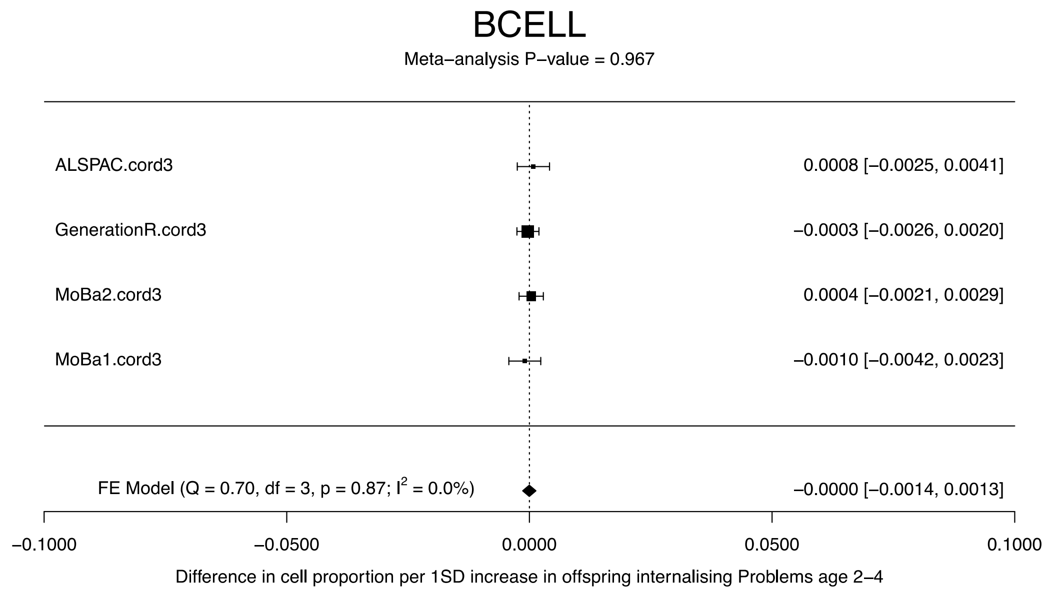


Z3 Leave-one-out plot of the childhood peripheral blood meta-analysis age 7 (Cg07283896)

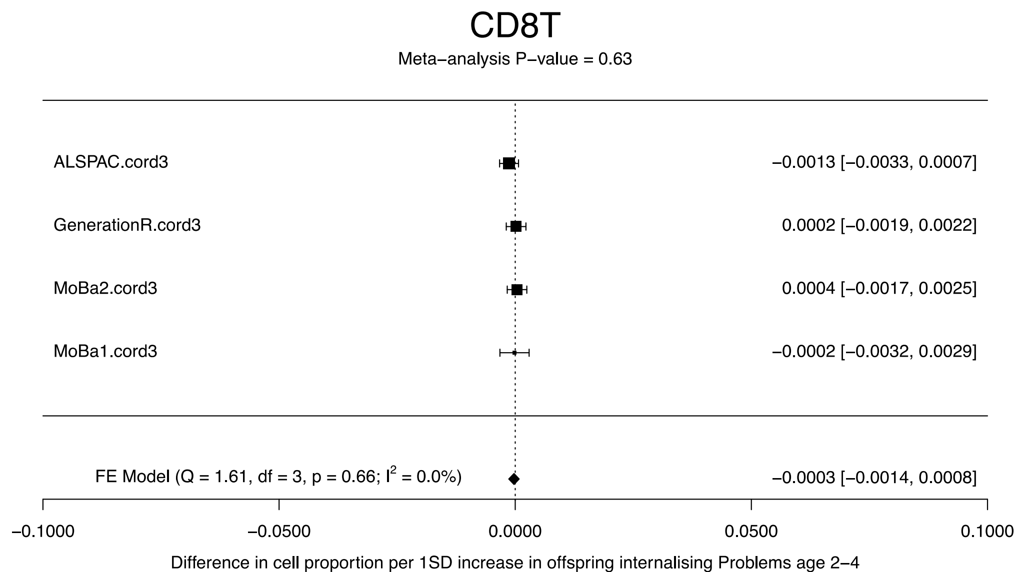


Appendix AA

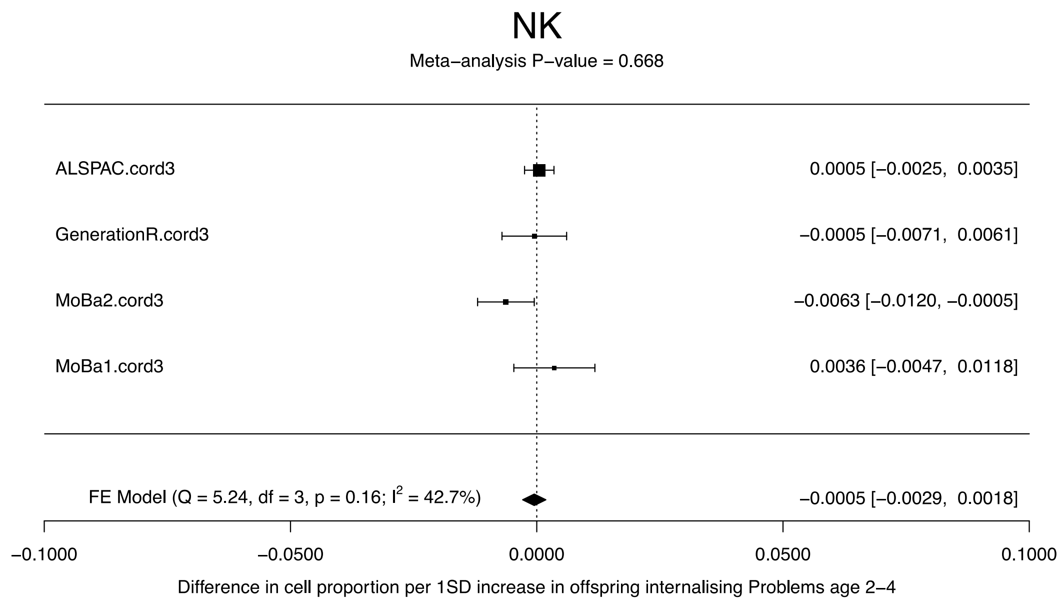
AA1 Forest plot of associations between cord blood BCELL cell proportion types and offspring internalising problems at age 3



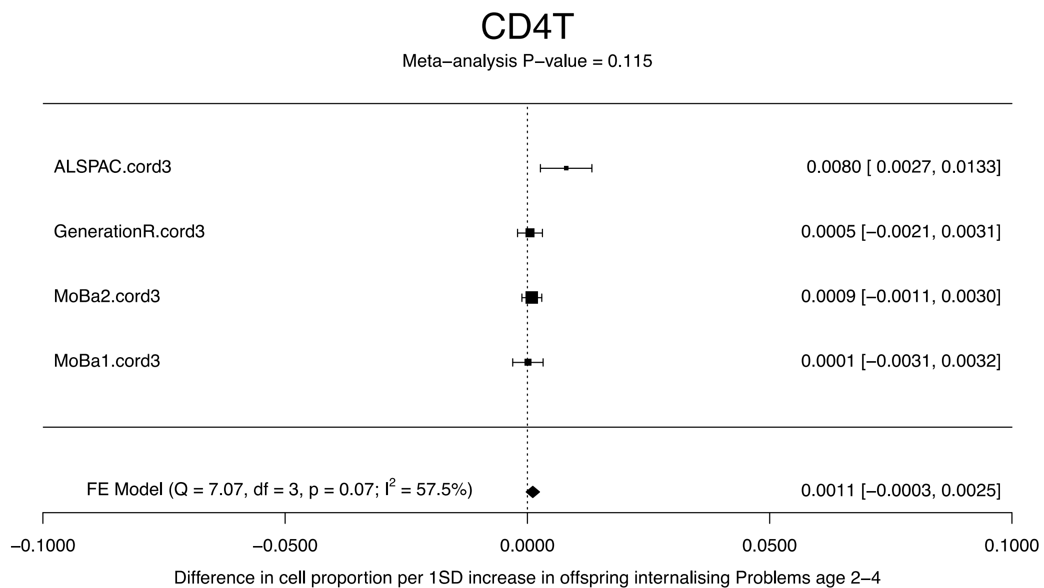
AA2 Forest plot of associations between cord blood CD8T cell proportion types and offspring internalising problems at age 3



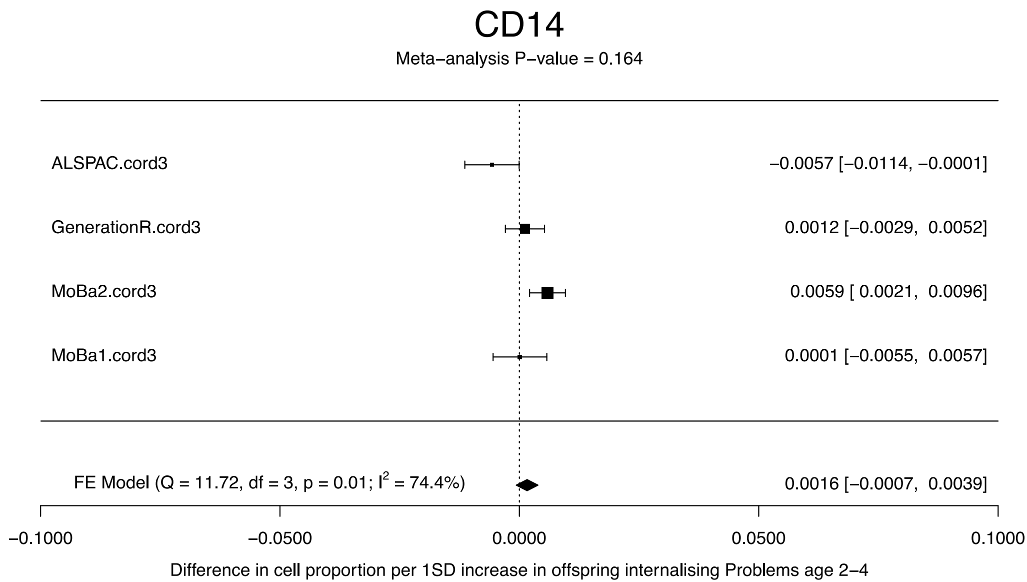
AA3 Forest plot of associations between cord blood NK cell proportion types and offspring internalising problems at age 3



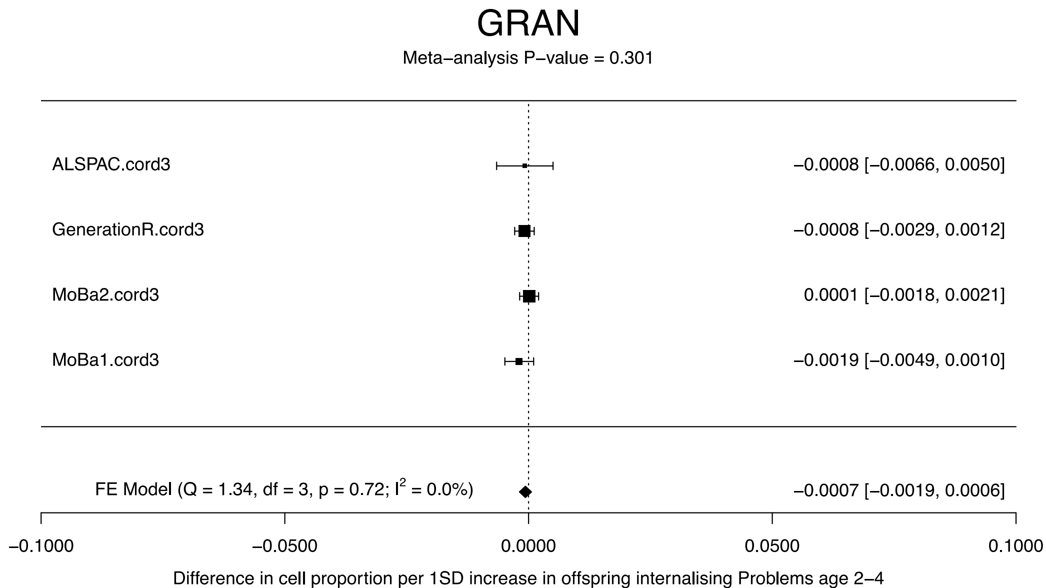
AA4 Forest plot of associations between cord blood CD4T cell proportion types and offspring internalising problems at age 3



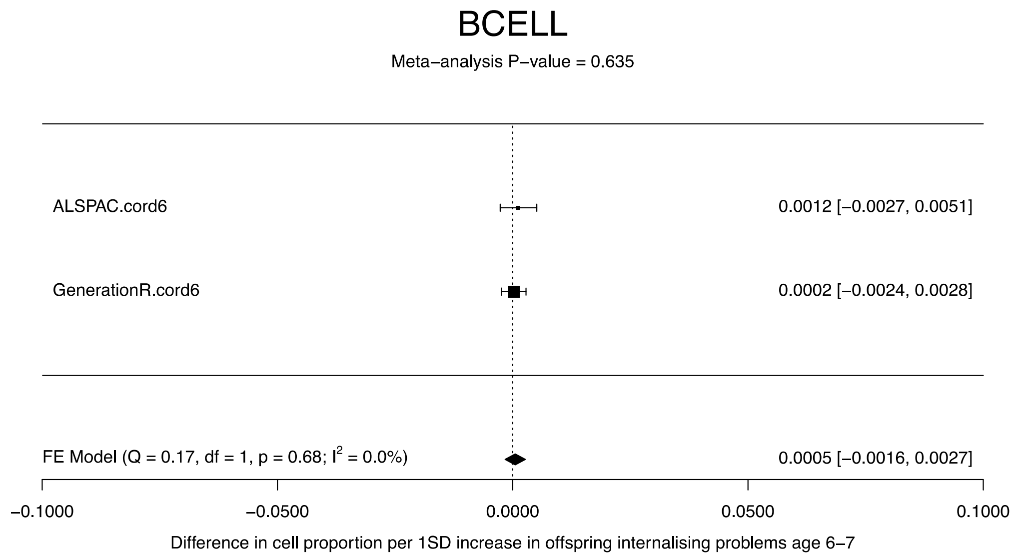
AA5 Forest plot of associations between cord blood CD14 cell proportion types and offspring internalising problems at age 3



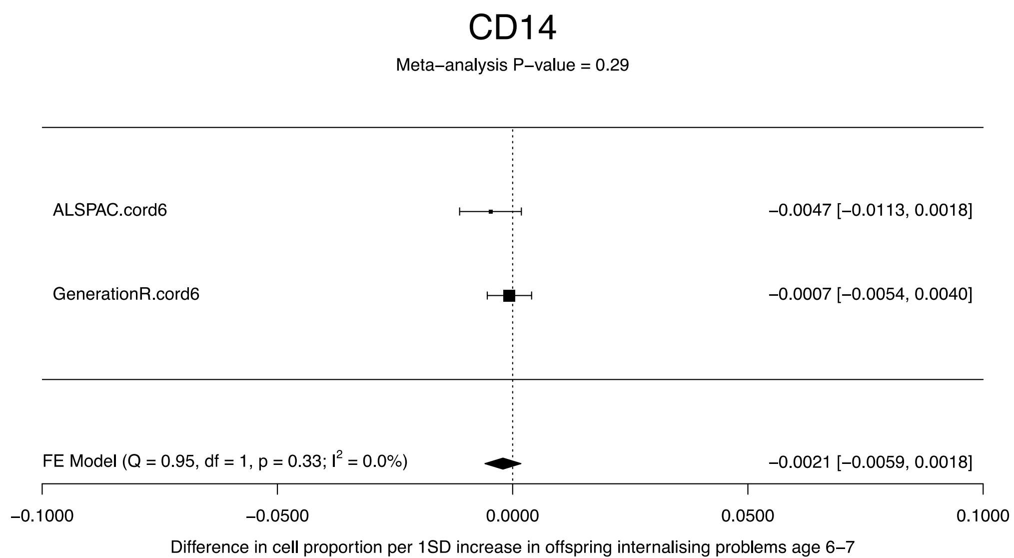
AA6 Forest plot of associations between cord blood GRAN cell proportion types and offspring internalising problems at age 3



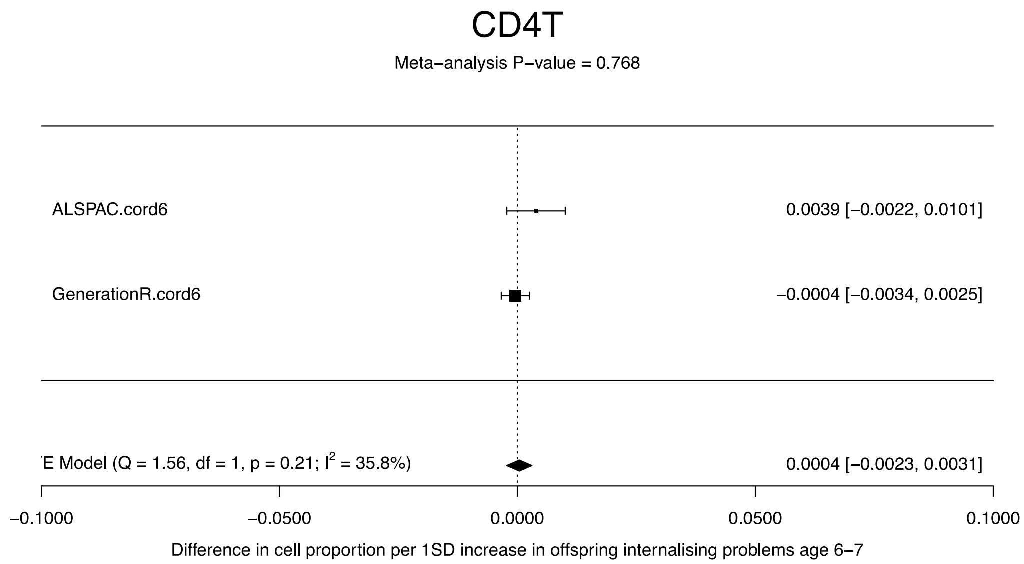
AA7 Forest plot of associations between cord blood BCELL cell proportion types and offspring internalising problems at age 7



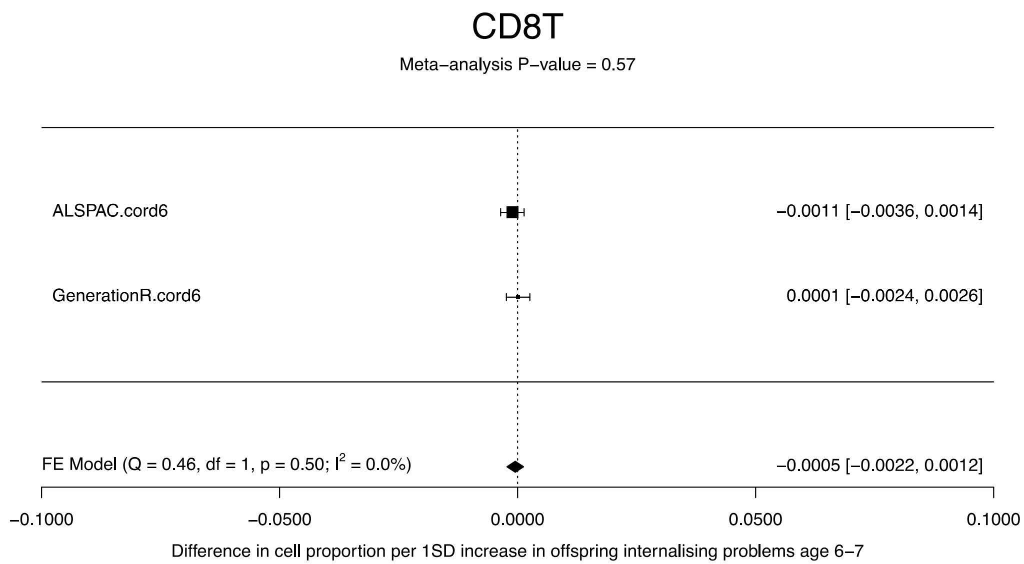
AA8 Forest plot of associations between cord blood CD14 cell proportion types and offspring internalising problems at age 7



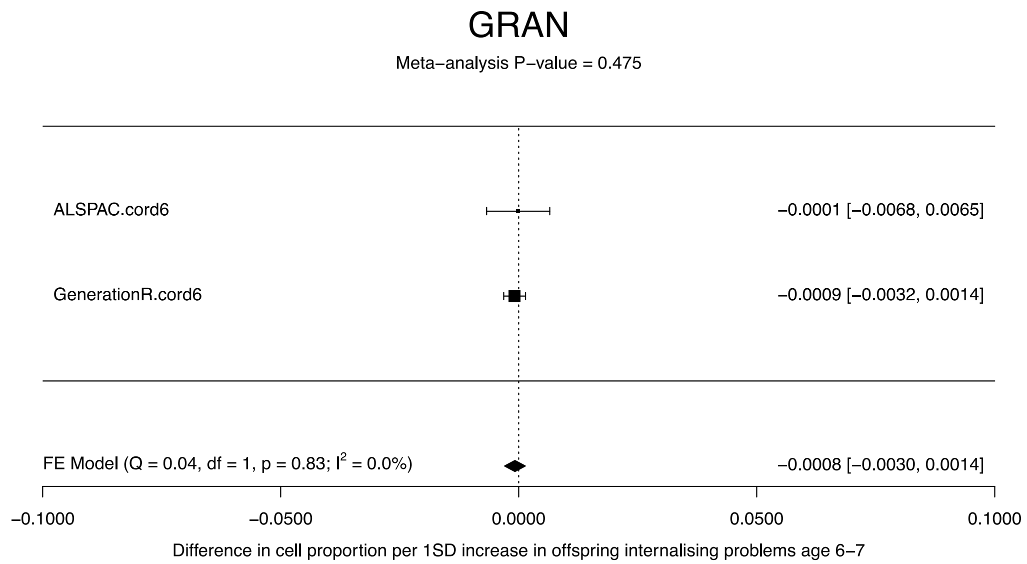
AA9 Forest plot of associations between cord blood CD4T cell proportion types and offspring internalising problems at age 7



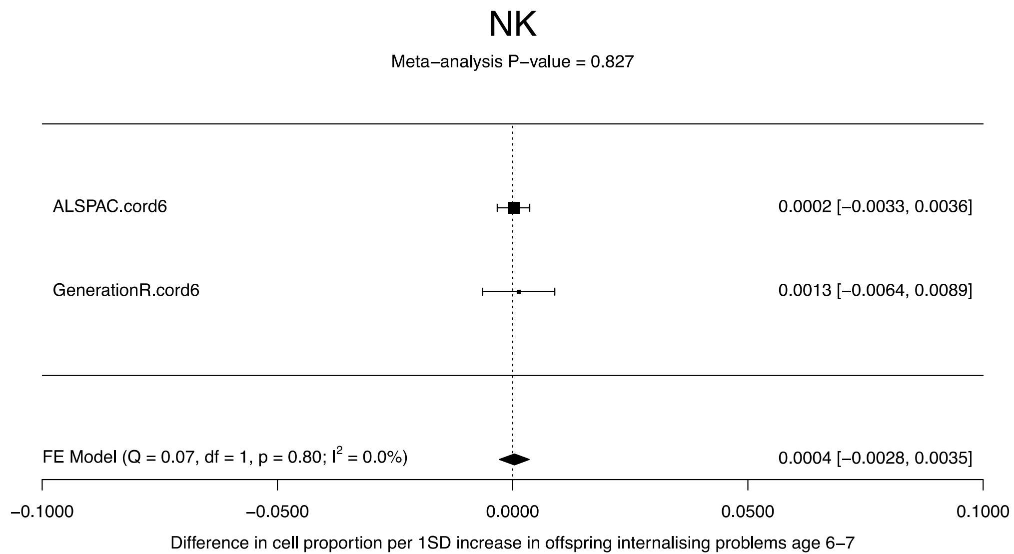
AA10 Forest plot of associations between cord blood CD8T cell proportion types and offspring internalising problems at age 7



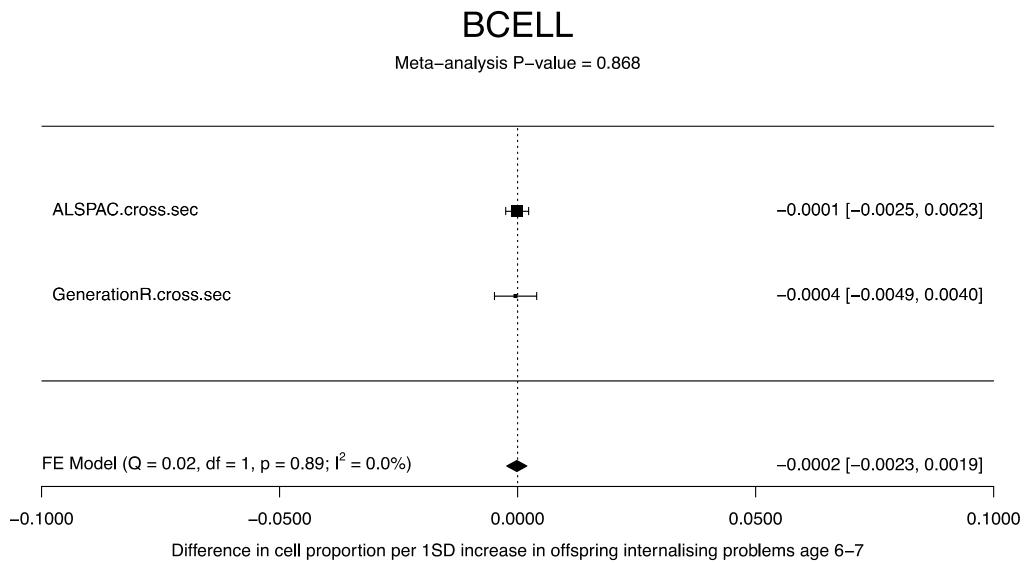
AA11 Forest plot of associations between cord blood GRAN cell proportion types and offspring internalising problems at age 7



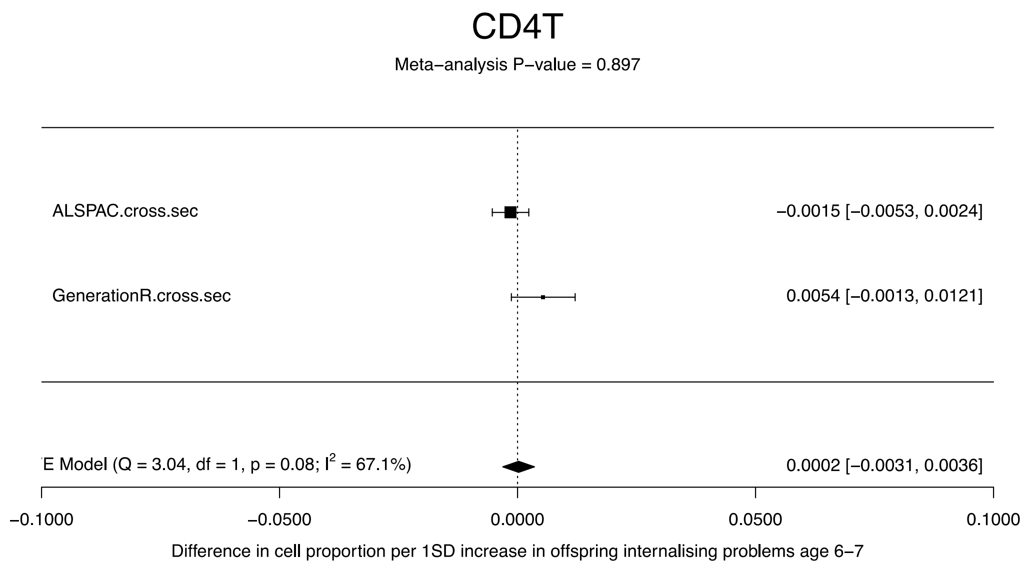
AA12 Forest plot of associations between cord blood NK cell proportion types and offspring internalising problems at age 7



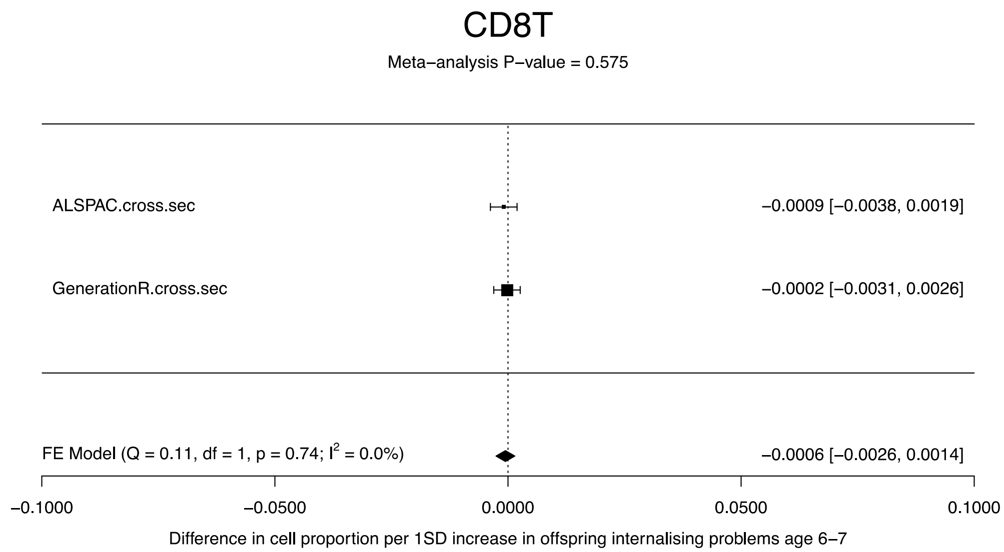
AA13 Forest plot of associations between peripheral blood BCELL cell proportion types and offspring internalising problems at the age of 7



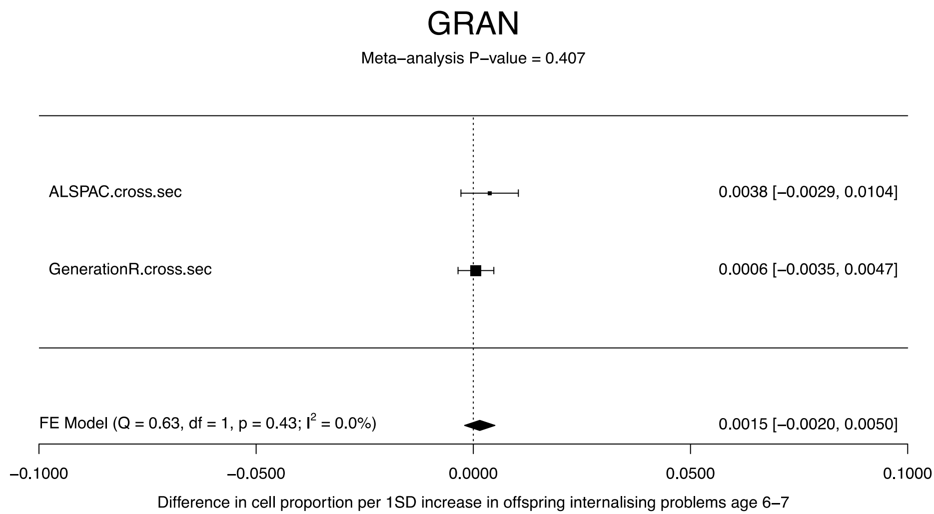
AA14 Forest plot of associations between peripheral blood CD4T cell proportion types and offspring internalising problems at age 7



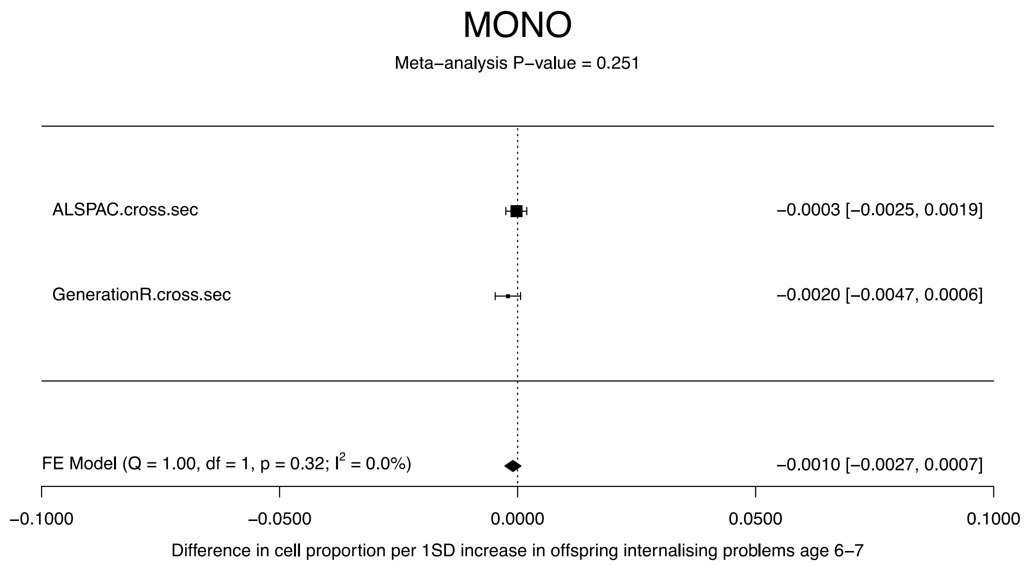
AA15 Forest plot of associations between peripheral blood CD8T cell proportion types and offspring internalising problems at age 7



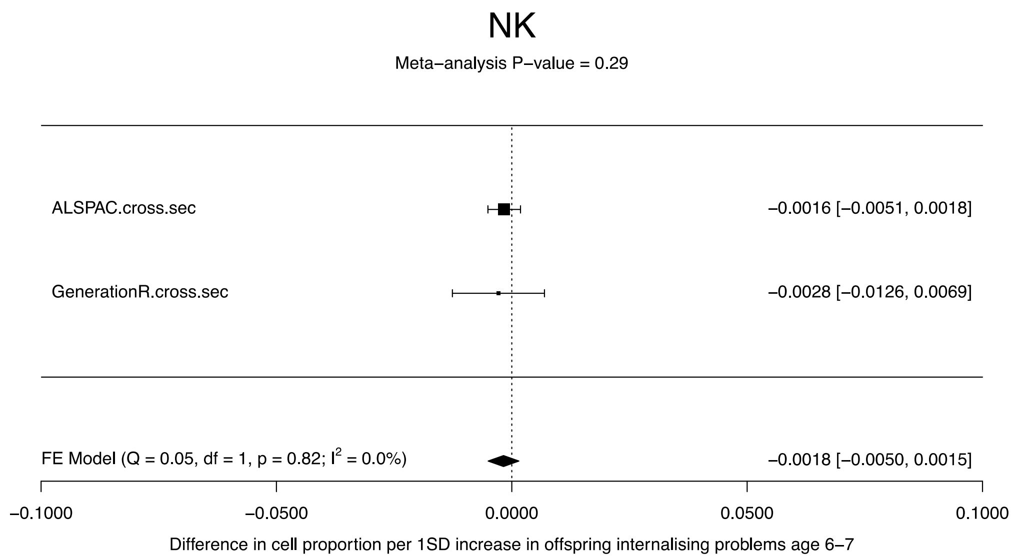
AA16 Forest plot of associations between peripheral blood GRAN cell proportion types and offspring internalising problems at age 7



AA17 Forest plot of associations between peripheral blood MONO cell proportion types and offspring internalising problems at age 7



AA18 Forest plot of associations between peripheral blood NK cell proportion types and offspring internalising problems at age 7



Appendix BB

BB1 Candidate gene-wide analyses

	Cord ~ Internalising problems age 3		Cord ~ Internalising problems age 7		Cross-sectional analysis	
CpG site (Gene)	**Estimate (SE)**	**P-Value**	**Estimate (SE)**	**P-Value**	**Estimate (SE)**	**P-Value**
Chr5:cg19457823 (NR3C1)	-1.17 (0.53)	0.027	0.15 (0.48)	0.763	0.44 (0.70)	0.529
Chr11: cg22043168 (BDNF)	4.66 (1.93)	0.016	1.65 (1.86)	0.374	-2.13 (2.13)	0.316
Chr11: cg13360150 (BDNF)	5.12 (2.22)	0.021	1.87 (1.99)	0.348	-5.92 (4.18)	0.157
Chr11: cg10022526 (BDNF)	-4.63 (2.01)	0.022	-4.24 (1.76)	0.016	-0.74 (2.58)	0.774
Chr11: cg22043168 (BDNF)	-1.43 (0.65)	0.028	0.33 (0.63)	0.601	2.19 (1.09)	0.045
Chr11: cg23497217 (BDNF)	-1.96 (0.91)	0.030	-0.97 (0.95)	0.305	0.20 (1.39)	0.887
Chr11: cg06025631 (BDNF)	-5.58 (2.74)	0.042	3.06 (3.44)	0.374	-10.31 (5.43)	0.058
Chr11: cg26949694 (BDNF)	-1.77 (0.87)	0.043	0.09 (0.78)	0.906	0.65 (1.31)	0.617
Chr11: cg15014679 (BDNF)	-1.05 (0.53)	0.048	-0.18 (0.52)	0.733	0.69 (0.68)	0.307
Chr6: cg07633853 (FKBP5)	-1.03 (0.39)	0.007	-0.27 (0.46)	0.554	0.28 (0.76)	0.716
Chr6: cg01294490 (FKBP5)	-2.94 (1.16)	0.011	-0.06 (1.17)	0.957	0.92 (1.74)	0.600
Chr6: cg14284211 (FKBP5)	-1.34 (0.56)	0.017	-0.70 (0.56)	0.216	0.12 (0.99)	0.906
Chr6: cg10300814 (FKBP5)	2.82 (1.29)	0.029	1.34 (1.22)	0.272	-0.45 (1.91)	0.815
Chr6: cg00130530 (FKBP5)	1.31 (0.601)	0.029	-0.17 (0.61)	0.776	-0.62 (0.82)	0.450

Appendix CC

CC1 Summary of results of top CpG sites from the probe-level analysis

Model	N CpG sites with $P < 1 \times 10^{-5}$	Range effect estimate
Internalising problems age 3 ~ Cord blood DNA methylation (N = 3,011)		
All offspring (minimally adjusted)*	2	2.50 to 3.07
All offspring (adjusted for covariates)	4	2.41 to 11.25
Female sex offspring (adjusted for covariates)	3	2.80 to 28.83
Male sex offspring (adjusted for covariates)	9	2.65 to 29.34
All offspring (adjusted for covariates and maternal anx/dep)	5	1.60 to 12.44
Internalising problems age 7 ~ Cord blood DNA methylation (N = 1,601)		
All offspring (minimally adjusted)*	4	1.10 to 3.84
All offspring (adjusted for covariates)	3	1.11 to 4.10
Female sex offspring (adjusted for covariates)	5	3.30 to 8.76
Male sex offspring (adjusted for covariates)	6	2.27 to 13.53
All offspring (adjusted for covariates and maternal anx/dep)	6	1.08 to 6.75
Internalising problems age 7 ~ Childhood DNA methylation age 7 (N = 1,121)		
All offspring (minimally adjusted)*	3	2.04 to 5.33
All offspring (adjusted for covariates)	2	2.54 to 5.35
Female sex offspring (adjusted for covariates)	2	23.84 to 72.60
Male sex offspring (adjusted for covariates)	7	3.25 to 8.59
All offspring (adjusted for covariates and maternal anx/dep)	2	2.87 to 5.47

*Note. * only adjusted for estimated cell counts and 20 surrogate variables. Covariates: maternal age, maternal smoking, maternal education, offspring age, estimated cell counts and 20 surrogate variables.*