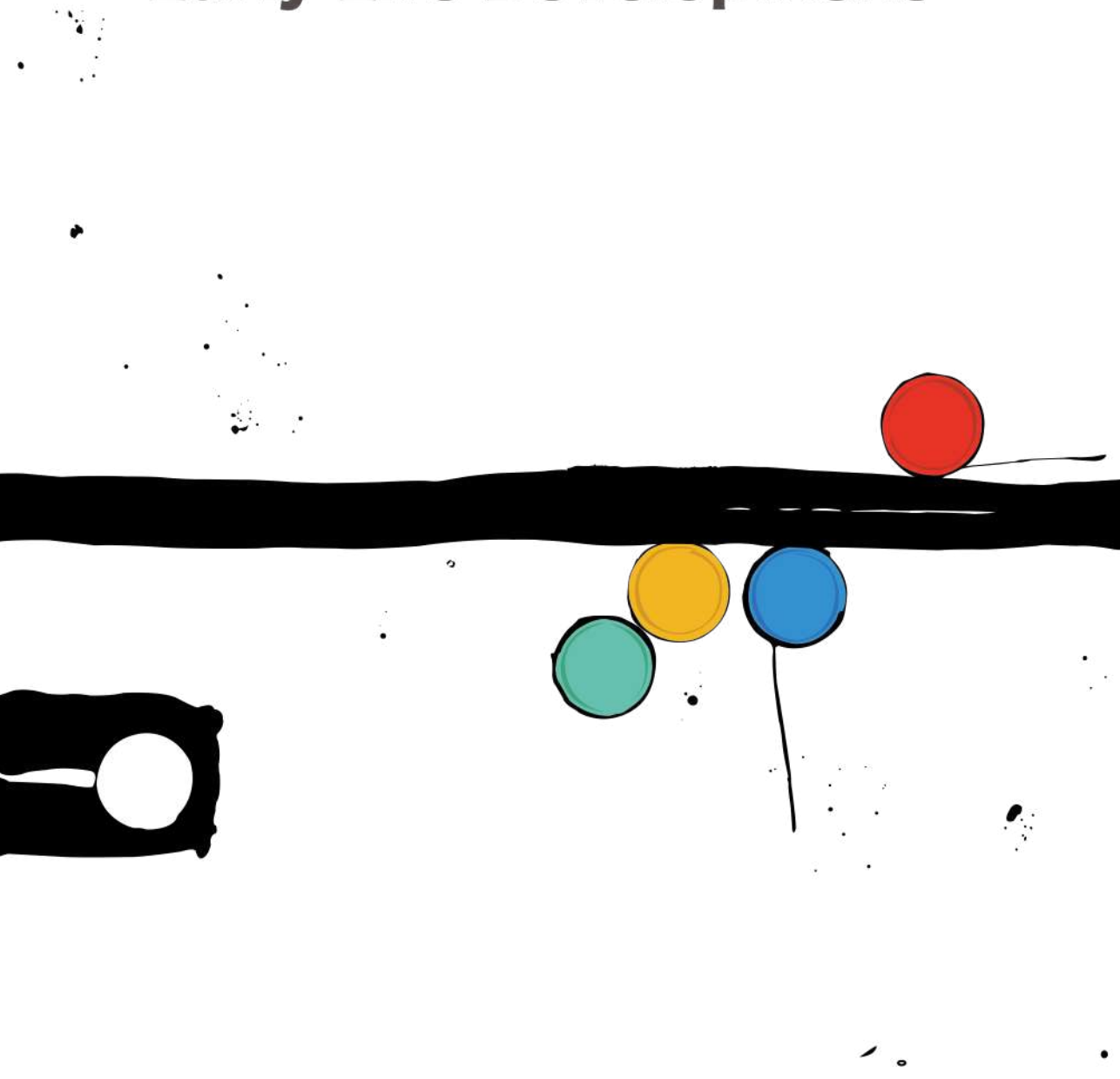
Genetics and epigenetics of early life development

Odintsova, Veronika

2022

document version
Publisher's PDF, also known as Version of record

Genetics and Epigenetics of Early Life Development



Veronika Odintsova

GENETICS AND EPIGENETICS OF EARLY LIFE DEVELOPMENT

Veronika Odintsova

paranymphs:
Camiel van der Laan
Floris Huider

Financial support for printing of this thesis was kindly provided by the Amsterdam Reproduction & Development (AR&D) research institute and European network staff eXchange for integrAting precision health in the health Care sysTems (ExACT) consortium (Horizon2020)

Design: John (Can) Lockman, inspired by the painting of Antonio Meneghetti “Cinque soli”, 1997, on the cover (with permission from Antonio Meneghetti Scientific Foundation)

Print: Proefschriftenprinten.nl

ISBN 978-90-832468-1-9

VRIJE UNIVERSITEIT

GENETICS AND EPIGENETICS OF EARLY LIFE DEVELOPMENT

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Gedrags- en Bewegingswetenschappen
op maandag 23 mei 2022 om 11.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Veronika Odintsova

geboren te Kazan, Rusland

promotoren:	prof.dr. D.I. Boomsma
	prof.dr. C.V. Dolan
copromotoren:	dr. J. van Dongen
	dr. C.E.M. van Beijsterveldt
promotiecommissie:	dr. E. van Bergen
	prof.dr. B.T. Heijmans
	dr. H. Mbarek
	prof.dr. T.J. Roseboom
	prof.dr. C. Schuengel
	dr. P.J.G. Zwijnenburg

Contents

ABBREVIATIONS

General

ALSPAC	Avon Longitudinal Study of Parents and Children
ANTR	Adult Netherlands Twin Register
BMI	body mass index
BMI	body mass index
chr	chromosome
CI	confidence interval
CpG	cytosine-phosphate-guanine
DSM	Diagnostic and Statistical Manual of Mental Disorders
DZ	dizygotic
EPIC	Infinium HumanMethylationEPIC BeadChip Kit (Illumina)
eQT	expression quantitative trait loci
EWAS	epigenome-wide association study
FDR	false discovery rate
GA	gestational age
GEE	generalized estimating equations
GTEx	genotype-tissue expression
GWAMA	genome-wide association meta-analysis
GWAS	genome-wide association study
GWS	genome-wide significant
GxE	gene-environment interaction
LD	linkage disequilibrium
mQTL	methylation quantitative trait locus
MS	methylation score
MZ	monozygotic
NTR	Netherlands Twin Register
OMIM	Online Mendelian Inheritance in Man
PBC	Peripheral blood cells
PGS	polygenic score
rGE	genome-environment correlation
SD	standard deviation
SE	standard error
SES	socio-economic status
SNP	single nucleotide polymorphism
YNTR	Young Netherlands Twin Register
450k	Infinium HumanMethylation450 BeadChip Kit (Illumina)

Chapters 5 and 6

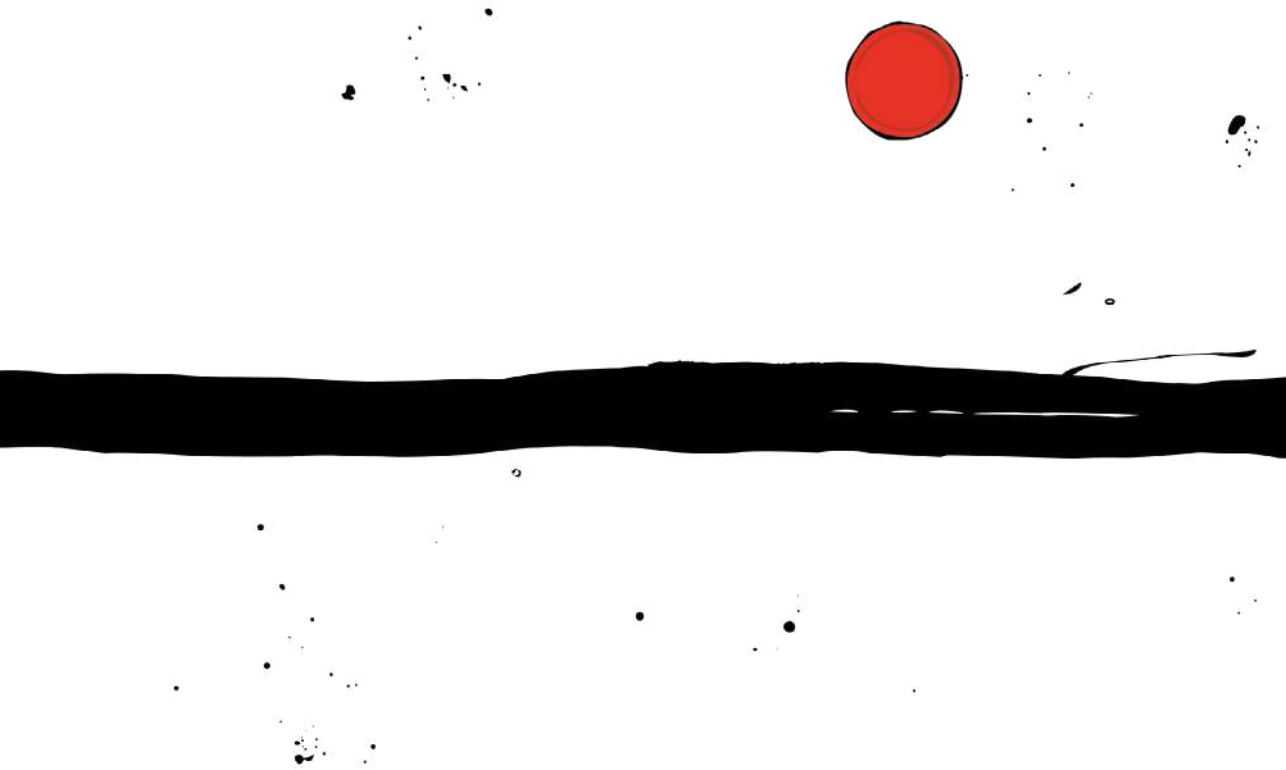
LH	left-handedness
MH	mixed-handedness
NRH	non-right-handedness
RH	right-handedness

Chapter 7

ASB	antisocial behavior
ASEBA	Achenbach System of Empirically Based Assessment scales
ASPD	antisocial personality disorder
BDHI	Buss-Durkee Hostility Inventory
BPAQ	Buss-Perry Aggression Questionnaire
CBCL	Child Behavioral Checklist
CD	conduct disorder
CGS	candidate gene studies
CPA	chronical physical aggression
CPRS-R:L	long version of the Conners Parent Rating Scale
CU	callous-unemotional traits
DP	dysregulation profile
ICD	International Classification of Diseases
iPSC	induced pluripotent stem cells
MPNI	Multidimensional Peer Nomination Inventory aggression scale
ODD	oppositional defiant disorder
PACS	Parental Account of Childhood Symptoms
PGQ	Multidimentional Personality Questionnaire
RDoC	Research Domain Criteria (National Institute of Mental Health)
SBQ	Social Behavior Questionnaire
SCID	Structured Clinical Interview for Disorders
SDQ	Strengths and Difficulties Questionnaire
SSTAS	Spielberger State-Trait Anger Scale
TWAS	transcriptome-wide association study

Chapter 1

Introduction



Every field of science goes through its own development cycle while it confronts the old and new puzzles, gets impulses from scientists, discovers breakthroughs and technologies that shift scientific paradigms[1]. In the past decades we have witnessed the lightning change of molecular biology, stimulated by advances in genome sequencing and cell biology. These changes have transformed multiple scientific fields, such as genetics, epidemiology, and medicine, and have generated new hybrid fields, such as pharmacogenomics and personalized medicine.

One development that changed human complex trait genetics is genome-wide association studies (GWAS). It quickly became evident that the search for genetic associations with complex traits and disease requires large sample sizes, whole-genome screens, and new analytical methods[2]. Many consortia, registers, and biobanks were stimulated by possibilities offered by GWAS, and generated an unprecedented amount of genetic and phenotypic data on different populations all over the globe, although still with an overrepresentation of European ancestries. As researchers, we made a huge step forward from the mindset of a single research study to a multi-project approach based on large-scale datasets. Importantly, developments in mega- and meta-analyses of multiple shared datasets and shared results have addressed issues of replication and statistical power. Increasingly, large population-based data address these research needs. These include the UK Biobank[3], FinnGen in Finland[4], Estonian Biobank[5], Netherlands and Swedish Twin registries[6,7]. In addition, researchers can obtain data by research-agreement-based access in collaboration with commercial parties, such as 23andMe[8].

From this perspective, twin registers are prominent research resources, with a unique longitudinal perspective. One of the first attempts of systematic data collection on twins was made in 1935 in the Soviet Union when the special twin kindergarten was created in the medical and genetic institute by initiative of Solomon Levit and colleagues, and followed 234 twins. The soviet psychologists and geneticists studied memory, development, attention, and intellect applying twin-control design[9]. This initiative was interrupted in 1938 in the period of Lysenkoism[10]. Consistently, twin registers have began appearing since the early 1950-ies[11] in Europe and then worldwide, with the aim of studying the role of genetics and environment for different phenotypes and diseases, and thus predated the genomic era. Today the international network of multiple twin registers all over the world contains information on many phenotypes and biosample collections obtained from twins and their families longitudinally. It covers almost all types of behavioral traits, mental health and various complex diseases, with total sample exceeding 1.3 million individuals[12].

The classical twin design is an elegant way to estimate the genetic and environmental influences on human traits. It is based on comparing the phenotypic resemblances of two types of twin pairs: monozygotic (one-zygote; MZ) twins and dizygotic (two-zygotes; DZ) twins. Developing from one zygote, monozygotic twins share almost all of their DNA sequence. Developing in parallel from two zygotes, dizygotic twins, like full siblings, share on average half of their segregating alleles.

Different patterns of correlations between phenotypes of MZ and DZ twins allow tests of hypotheses regarding the contribution of genetic and environmental factors to phenotypic individual differences. The first hypothesis is that individual differences in the phenotype

of interest are due to non-genetic factors, i.e. shared and unshared environmental influences. This implies that the correlations for monozygotic and dizygotic twin pairs should be equal (and greater than zero if there is an influence of shared environment). The second hypothesis is that individual differences in the phenotype of interest are explained by genetic variation. This implies that the correlations for monozygotic twins are higher than for dizygotic twins. Boomsma et al. wrote that applications of twin analysis had led to a change in our understanding of the determinants of health and disease[13] and pointed to the value of data in monozygotic twins: "The concordance between monozygotic twins sets the upper limit on predictions of the human genome sequence"[13]. As noted by Martin et al., twin studies re-oriented research programs suggesting higher heritability in some cases, when the search was focused on environmental factors, and lower heritability in other cases, when the genetics was considered the main cause of a trait[14]. The comparison of monozygotic and dizygotic twin resemblances clearly showed that genetics contributes significantly to the variation in virtually all human phenotypes[15]. The twin design gives opportunities to meet challenges in studying new traits and environmental changes such as impact of human-driven environmental change and the COVID-19 pandemic on health[16], different roles of men and women in society[17], impact of digital education on child cognitive and socio-emotional development[18–20], effect of spaceflights on human body[21] and also the high-dimensional data generated by omics technologies.

Responding to the development of omics technologies (e.g., high-throughput molecular profiling of genes, DNA methylation, transcripts, proteins, and metabolites, a growing number of twin registries have collected molecular data, including genome, exome, epigenome, metabolome, and microbiome[22]), thereby providing new insights into the biological processes that underlie the association between genes and complex phenotypes and assessing the extent to which their variation is heritable.

A major focus of this thesis is beyond the variation in the DNA sequence, at the level of the epigenome. Epigenetic regulation is attributable to specific mechanisms which control the functional activity of gene promoters and other regulatory regions of the genome and include DNA methylation, histone modifications, chromatin remodeling and RNA-based mechanisms. DNA methylation is an epigenetic mark which regulates gene expression through adding a methyl group (CH_3) on cytosine-phosphate-guanine (CpG) dinucleotides in the genome. DNA methylation arrays provide a tool to assess whole-epigenome profiles with coverage of up to 850,000 CpGs on high-throughput arrays and are widely used in large-scale epigenetic studies.

An epigenome-wide association study (EWAS) incorporates a design that is used to detect associations between DNA methylation levels at the loci across the genome and traits in samples from populations. EWASs have the potential to bridge the knowledge gap from genome sequence to a trait or disease. While the genetic architecture comprises the number of variants influencing a phenotype, the magnitude of their effects on the phenotype, the population frequency of these variants and their interactions with each other and the environment[23], epigenetics adds an important layer of information on how the genome is packed and expressed[24]. As the epigenetic mechanisms themselves are influenced by genome and environment, epigenetics is a kind of "dynamic interface"

that reflects the information partly from the genome and partly from environmental influences. As shown by van Dongen et al., monozygotic twins can show high similarity of methylation across genome-wide CpGs, but the slight difference of variance in DNA methylation indicates a certain degree of stochastic and environmental influences[25]. Is DNA methylation a source of phenotypic variance? If so, classical twin modeling can be extended by taking into account the epigenetic effects that can be added as an independent third source of variance, distinguishable from the variance attributable to unshared environment, or from both genetic and environmental components[26–28].

With respect to phenotyping, a major focus of the thesis is on early life development. Notwithstanding a century of twin reseach on complex traits[13,14,22], early life traits and characteristics have been somewhat neglected. Polderman and colleagues[15] meta-analyzed heritability estimates based on publications that used the classical twin design. The meta-analysis across all traits included more than 14 million twin pairs. The analysis of the traits evaluated in twin studies showed a dominant focus on traits from the psychiatric, metabolic and cognitive domains. Only ~0.6% of traits and ~1.3% of studies were related to the perinatal and early life period, or to disorders with onset during childhood. Traits related to pregnancy, childbirth, neonatal and early-life periods, except congenital malformations and chromosomal abnormalities, tended to have low heritability (h^2 = 0.117, s.e. = 0.069), and were influenced by shared environmental effects, including for example maternal effects (c^2 = 0.310, se = 0.033) to a larger degree than traits that expressed later in life.

The presence of genetic effects has inspired genetic association studies of early life traits (or traits supposed to originate in early period, such as handedness) with pre- and perinatal characteristics, with birth weight as an example of a trait for which many variants now have been found[29].

There are now possibilities to search for factors beyond the genetic sequence. The prenatal and early discordance in monozygotic twins is unlikely to be caused by the DNA sequence, although discordant MZ twins have inspired sequencing studies to search for post-twinning mutations[30], but such mutations seem rare. Therefore, the search for explanations for MZ discordance has included a variety of ante-natal epigenetic and environmental influences[31,32].

The concepts of the developmental origins of health and disease (DOHaD), the fetal origins hypothesis, fetal programming, prenatal programming, and the thrifty phenotypes hypothesis[33–36] posit that the development of disease later in life is caused in part by the influence of prenatal environmental factors on the fetus, rather than by genetic factors. The hypothesis of fetal programming was supported by prominent famine studies that analysed health outcomes in individuals, who, prenatally, suffered prolonged exposure to famine. During World War II, there was a severe food limitation under occupation that spanned a period of 5 months in the Western part of the Netherlands, the so-called Dutch Hunger Winter. According to the Dutch Hunger Winter Study, the exposure to famine during all stages of the prenatal period was associated with reduced glucose tolerance, an increased risk of type 2 diabetes, and during late stage of the prenatal period with obesity and increased rates of cardiovascular disease[37]. According to the Leningrad

Siege Study, which address the consequences of the 28-month siege of Leningrad, the intrauterine exposed group evidenced endothelial dysfunction by higher concentrations of von Willebrand factor and a stronger interaction between adult obesity and blood pressure[38]. Both studies demonstrated the long-term effects of intrauterine malnutrition on health. Epigenetic mechanisms were proposed as one of explanation[37]. According to the Dutch Hunger Winter Study, maternal famine in early gestation has been associated with long-term DNA methylation changes that were still observable in adulthood[39,40].

Likewise, the epigenetic studies of early life exposures provide more evidence that DNA methylation could mediate their effect on later outcome. A large meta-analysis of prenatal maternal smoking indicates that associated DNA methylation changes are to some extent stable throughout the lifespan and are located at genes related to vast variety of disease and neurodevelopmental disorders[41]. All and all, the epigenetic mechanisms are an important element in understanding the developmental origins of later life disease risk[42–44].

Increasingly, genetics and epigenetics are moving towards prediction at the individual level, with construction of polygenic scores based on results from GWAS, and DNA methylation scores based on results from EWAS[45]. The open question remains how much variance in a complex trait or exposure can be explained by DNA methylation scores, and how it can be combined with polygenic score prediction.

OUTLINE

My thesis builds on approaches, which are leading in contemporary genetic epidemiology and have been stimulated by molecular biology advancements, and applies these approaches to human traits mainly from the first years of life. In several chapters, I first conducted an extensive search of the literature. To this end, I searched the literature using a machine learning and systematic approach for analysis of published associations. Then I performed epidemiological studies to identify statistical associations between early life characteristics and Apgar scores and handedness. The epigenome-wide association study in this thesis was realized through a search of differentially methylated positions and differentially methylated regions in association with a trait or exposure, through comparison of discordant monozygotic twins, through longitudinal analysis of DNA methylation, and through DNA methylation score prediction of traits or exposures. I have benefitted from data collected in the Netherlands Twin Register (NTR) and Avon Longitudinal Study of Parents and Children (ALSPAC). These registers collected data from newborns and children to address questions regarding the etiology of individual differences in developmental traits.

The first part **"Early Life Outcomes and Exposures"** is focused on early life development by reviewing the literature and by association studies using data of the Netherlands Twin Register. **Chapter 2** is dedicated to one of the earliest newborn measurements at birth, i.e. Apgar scores. It investigates whether Apgar scores, an indicator of newborn's state at first minutes after birth that consists of five criteria (**A**ppearance, **P**ulse, **G**rimace, **A**ctivity, and **R**espiration), are heritable, and which prenatal and perinatal characteristics effect

Apgar scores in twins. **Chapter 3** introduces DNA methylation as one of the epigenetic mechanisms and describes the DNA methylation changes in children exposed to maternal smoking during pregnancy. **Chapter 4** investigates DNA methylation differences in children who were breastfed and who were not breastfed. **Chapter 5** searches for associations between handedness – a phenotype was shown by previous studies to appear in prenatal period – and early life characteristics. The association between left-handedness and DNA methylation in different tissues and ages is analyzed in **Chapter 6**.

The second part of my thesis **"Genetic and Epigenetic Prediction of Complex Traits"** takes a broader view on several complex traits and exposures through the life span. **Chapter 7** demonstrates the application of the machine learning based selection of literature for systematic review and analysis of large body of literature and applies this approach to genetics and epigenetics of aggression and provides a comprehensive overview of the field of GWASs, EWASs and omics studies of aggression. **Chapter 8** investigates predictions based on polygenic scores and DNA methylation profiles in different tissues for birth weight, prenatal maternal smoking, body mass index (BMI), former and current smoking status.

The third part **"Twin Design for Biomedical Research: Methodological issues"** covers several methodological aspects important for conducting population-based twin research. **Chapter 9** describes in detail the establishment of twin register as a resource for epidemiological, genetic, epigenetic, and biomarker studies. **Chapter 10** discusses the question whether twin research findings can be extrapolated to singletons and which specific differences are related to twinning and describes a large-scale comparison of twins and singletons through the life span. **Chapter 11** is devoted to a consideration of ethical issues of twins in research.

References

1. Kuhn, T. S. The structure of scientific revolutions. In: International Encyclopedia of Unified Science (University of Chicago Press, 1970).
2. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *American journal of human genetics* **101**, 5–22 (2017).
3. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *American journal of epidemiology* **186**, 1026–1034 (2017).
4. FinnGen research project is an expedition to the frontier of genomics and medicine | FinnGen. https://www.finngen.fi/en.
5. The Estonian Biobank - EIT Health Scandinavia. https://www.eithealth-scandinavia.eu/biobanks/the-estonian-biobank/.
6. Nederlands Tweelingen Register | Nederlands Tweelingen Register. https://tweelingenregister.vu.nl/.
7. The Swedish Twin Registry | Karolinska Institutet. https://ki.se/en/research/the-swedish-twin-registry.
8. 23andMe Research Innovation Collaborations Program. https://research.23andme.com/research-innovation-collaborations/.
9. Levit, S. G. Twin Investigations in the U. S. S. R. *Journal of Personality* **3**, 188–193 (1935).
10. Ptushenko, V. v. The pushback against state interference in science: how Lysenkoism tried to suppress Genetics and how it was eventually defeated. *Genetics* **219**, iyab162 (2021).
11. Hur, Y.-M. & Craig, J. M. Twin registries worldwide: an important resource for scientific research. *Twin research and human genetics: the official journal of the International Society for Twin Studies* vol. 16 1–12 (2013).
12. Hur, Y., Odintsova, V., Ordonana, J., Silventoinen, K. & Willemsen, G. Twin Registries Worldwide. in *Twin and Family Studies of Epigenetics* (eds. Harris, J., Segal, N., Tarnoki, A. D. & Tarnoki, D. L.) (Elsevier, 2022).
13. Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. Nature reviews. *Genetics* **3**, 872–882 (2002).
14. Martin, N., Boomsma, D. & Machin, G. A twin-pronged attack on complex traits. *Nature genetics* **17**, 387–392 (1997).
15. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature genetics* **47**, 702–709 (2015).
16. Williams, F. M. K. *et al.* Self-Reported Symptoms of COVID-19, Including Symptoms Most Predictive of SARS-CoV-2 Infection, Are Heritable. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **23**, 316–321 (2020).
17. Polderman, T. J. C. *et al.* The Biological Contributions to Gender Identity and Gender Diversity: Bringing Data to the Table. *Behavior genetics* **48**, 95–108 (2018).
18. Long, E. C. *et al.* The Genetic and Environmental Contributions to Internet Use and Associations With Psychopathology: A Twin Study. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **19**, 1–9 (2016).
19. Von Stumm, S. & Plomin, R. Monozygotic twin differences in school performance are stable and systematic. *Developmental science* **21**, e12694–e12694 (2018).
20. Li, M., Chen, J., Li, N. & Li, X. A Twin Study of Problematic Internet Use: Its Heritability and Genetic Association With Effortful Control. *Twin Research and Human Genetics* **17**, 279–287 (2014).
21. Garrett-Bakelman, F. E. *et al.* The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. *Science* **364**, (2019).
22. Van Dongen, J., Slagboom, P. E., Draisma, H. H. M., Martin, N. G. & Boomsma, D. I. The continuing value of twin studies in the omics era. *Nature Reviews Genetics* **13**, 640–653 (2012).

23. Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J. & Richards, J. B. Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nature reviews. Genetics* **19**, 110–124 (2018).
24. Gibney, E. R. & Nolan, C. M. Epigenetics and gene expression. *Heredity* **105**, 4–13 (2010).
25. Van Dongen, J. *et al.* Epigenetic variation in monozygotic twins: a genome-wide analysis of DNA methylation in buccal cells. *Genes* **5**, 347–365 (2014).
26. Dolan, C., Nivard, M. G., van Dongen, J., van der Sluis, S. & Boomsma, D. Advances in Genomics and Genetics Dovepress Methylation as an epigenetic source of random genetic effects in the classical twin design. *Advances in Genomics and Genetics* 5–305 (2015) doi:10.2147/AGG.S46909.
27. Bell, J. T. & Spector, T. D. A twin approach to unraveling epigenetics. *Trends in genetics: TIG* **27**, 116–125 (2011).
28. Molenaar, P. C., Boomsma, D. I. & Dolan, C. v. A third source of developmental differences. *Behavior genetics* **23**, 519–524 (1993).
29. Beck, J. J. *et al.* Genetic meta-analysis of twin birth weight shows high genetic correlation with singleton birth weight. *Human Molecular Genetics* (2021).
30. Zwijnenburg, P. J. G., Meijers-Heijboer, H. & Boomsma, D. I. Identical but not the same: The value of discordant monozygotic twins in genetic research. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **153B**, n/a-n/a (2010).
31. Castillo-Fernandez, J. E., Spector, T. D. & Bell, J. T. Epigenetics of discordant monozygotic twins: implications for disease. *Genome Medicine* **6**, 60 (2014).
32. Oates, N. A. *et al.* Increased DNA methylation at the AXIN1 gene in a monozygotic twin from a pair discordant for a caudal duplication anomaly. *American journal of human genetics* **79**, 155–162 (2006).
33. Barker, D. J. Fetal origins of coronary heart disease. *BMJ (Clinical research ed.)* **311**, 171–174 (1995).
34. Bianco-Miotto, T., Craig, J. M., Gasser, Y. P., van Dijk, S. J. & Ozanne, S. E. Epigenetics and DOHaD: from basics to birth and beyond. *Journal of developmental origins of health and disease* **8**, 513–519 (2017).
35. Godfrey, K. M. & Barker, D. J. Fetal programming and adult health. *Public health nutrition* **4**, 611–624 (2001).
36. Vaag, A. A., Grunnet, L. G., Arora, G. P. & Brøns, C. The thrifty phenotype hypothesis revisited. *Diabetologia* **55**, 2085–2088 (2012).
37. Roseboom, T. J. Epidemiological evidence for the developmental origins of health and disease: effects of prenatal undernutrition in humans. *Journal of Endocrinology* **242**, T135–T144 (2019).
38. Stanner, S. A. & Yudkin, J. S. Fetal programming and the Leningrad Siege study. *Twin research: the official journal of the International Society for Twin Studies* **4**, 287–292 (2001).
39. Heijmans, B. T. *et al.* Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17046–9 (2008).
40. Tobi, E. W. *et al.* DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nature Communications* **5**, 5592 (2014).
41. Joubert, B. R. *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *The American Journal of Human Genetics* **98**, 680–696 (2016).
42. Wadhwa, P., Buss, C., Entringer, S. & Swanson, J. Developmental Origins of Health and Disease: Brief History of the Approach and Current Focus on Epigenetic Mechanisms. *Seminars in Reproductive Medicine* **27**, 358–368 (2009).

43. Gluckman, P. D., Hanson, M. A. & Beedle, A. S. Early life events and their consequences for later disease: A life history and evolutionary perspective. *American Journal of Human Biology* **19**, 1–19 (2007).

44. Bateson, P. *et al.* Developmental plasticity and human health. *Nature* **430**, 419–421 (2004).

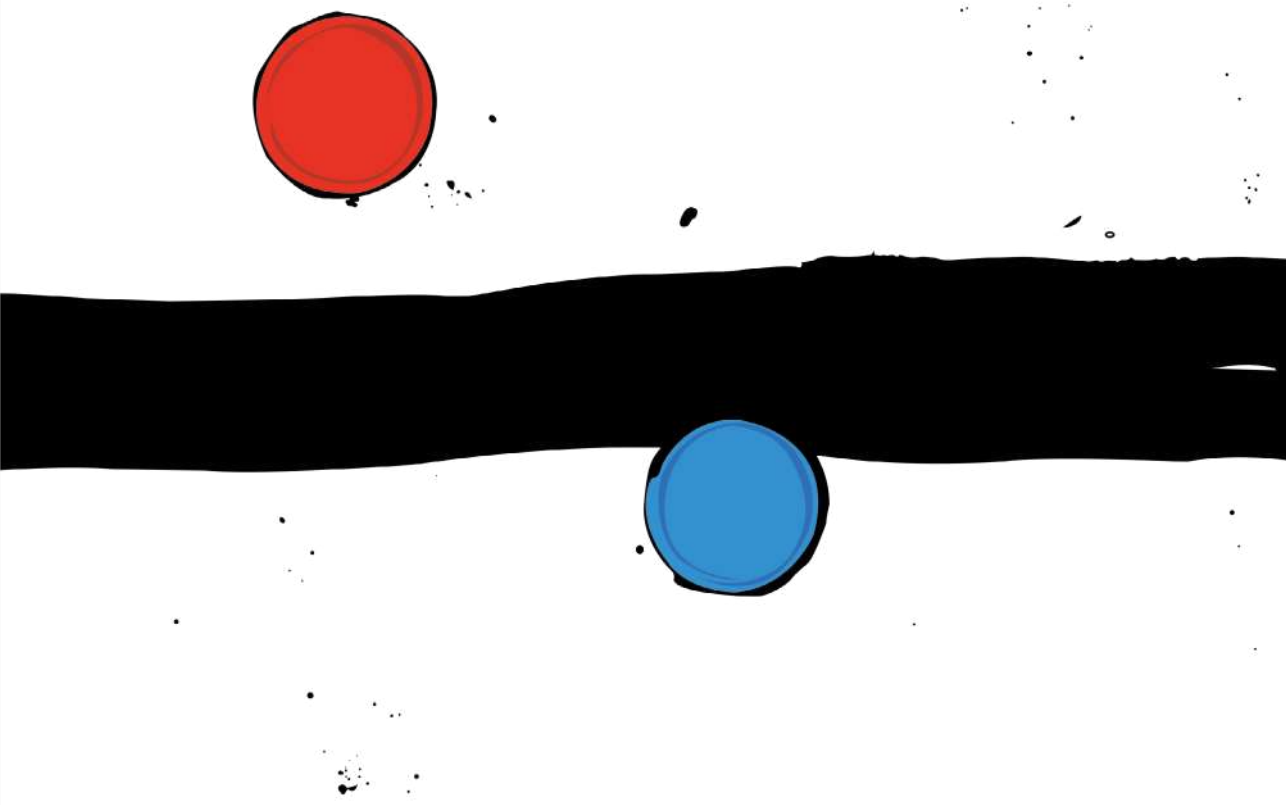
45. Shah, S. *et al.* Improving Phenotypic Prediction by Combining Genetic and Epigenetic Associations. *The American Journal of Human Genetics* **97**, 75–85 (2015).

PART I

Early life outcomes and exposures

Chapter 2

Pre- and perinatal characteristics associated with Apgar scores in a review and in a new study of Dutch twins



Abstract

A literature review was carried out to identify pre- and perinatal characteristics associated with variation in Apgar scores in population-based studies. The parameters identified in the literature search were included in the classical twin design study to estimate the effects of pre- and perinatal factors, both shared and non-shared by twins and to test for a contribution of genetic factors in 1- and 5-min Apgar scores in a large sample of Dutch monozygotic (MZ) and dizygotic (DZ) twins. The sample included MZ and DZ twins (N=5181 pairs) recruited by the Netherlands Twin Register (NTR) shortly after birth, with data on prenatal characteristics and Apgar scores at first and/or fifth minutes. Ordinal regression and structural equation modelling were used to analyse the effects of characteristics identified in the literature review and to estimate genetic and non-genetic variance components. The literature review identified 63 papers. Consistent with the review, we observed statistically significant effects of birth order, zygosity, and gestational age (GA) for 1- and 5-min Apgar scores of both twins. Apgar scores are higher in first-born vs second-born twins, and in dizygotic first-born vs monozygotic first-born twins. Birth weight had an effect on the 5-min Apgar of the first-born. Fetal presentation and mode of delivery had different effects on Apgar scores of first- and second-born twins. Parental characteristics and chorionicity did not have significant main effects on Apgar scores. The MZ twins' Apgar correlations equalled the DZ Apgar correlations. Our analyses suggest that individual differences in 1- and 5-min Apgar scores are attributable to shared and non-shared pre- and perinatal factors, but not to genotypic factors of the newborns. The main predictors of Apgar scores are birth order, zygosity, gestational age, birth weight, mode of delivery, and fetal presentation.

Published as: Odintsova, V. V., Dolan, C. V., van Beijsterveldt, C., de Zeeuw, E. L., van Dongen, J., & Boomsma, D. I. (2019). Pre- and perinatal characteristics associated with Apgar scores in a review and in a new study of Dutch twins. *Twin Research and Human Genetics,* 22(3): 164–176.

Introduction

The Apgar score is an important indicator of a newborn's health, and is established immediately after birth[1]. Apgar scoring is standard in obstetrics and neonatological practice, and the advantage of this screening tool is that it allows for prompt standardized assessment[1,2]. It has proven its utility as a population-level indicator of outcome risk, with predictive value for neonatal and infant death and post-neonatal development[3–6]. Its value has been confirmed in national register studies[7–12], and is recognized by national guidelines and the World Health Organization[13,14]. However, there is a discussion regarding other, possibly more precise, monitoring tools[15,16]. Apgar scores are evaluated at the first and fifth minute after birth.While a low 1-min score is transitory, persistence of poor health resulting in a low 5-min score often implies complications of clinical importance, and indicates that the newborn has not responded to initial intervention[4,5,12,14,17].

The assessment of Apgar scores in twins was first presented by Virginia Apgar in 1953. For twins, the Apgar score of the first-born twin was typically better than that of the second-born twin[1]. Replication studies of differences among first-and second-born twins[18–21], among twins and singletons[12,22], among term and preterm new-borns[20,23], and among different presentation and modes of delivery[24] followed this initial work. Two small studies[25,26] suggested that genetic factors also contribute to variation in Apgar scores. However, because of small sample sizes the results of these studies were inconclusive.

The aim of the present study is twofold. First, we present a literature review of the studies on Apgar scores, including singleton and multiple births, in order to identify pre- and perinatal characteristics associated with variation in Apgar scores in population-based studies. Second, in a large sample of Dutch monozygotic (MZ) and dizygotic (DZ) twins enrolled in the Netherlands Twin Register[27] we estimated effects of pre- and perinatal factors shared and non-shared by twins, including the parameters identified in the literature search and test for a contribution of genetic factors in the classical twin design.

Methods

Literature Review

A review of the literature regarding pre- and perinatal characteristics associated with Apgar scores was conducted in PubMed (MEDLINE), Web of Science, Embase and reference lists of retrieved articles. Search terms were "Apgar scores" and "heritability", "genetic effect", "prenatal factors", "twins", "fetal presentation", "mode of delivery", "gestational age", and "neonatal outcome". Studies on specific clinical aspects of pregnancy and neonatology, and on mortality and morbidity were excluded. We followed prior research on Apgar scores[11,28,29] and grouped pre- and perinatal characteristics in the following categories – biological maternal and paternal factors, socioeconomic factors, mode of conception, gestational age (GA), pregnancy and delivery characteristics, and newborn characteristics. The list of characteristics was included in an empirical study in which resemblance in mono- and dizygotic twins in a bivariate (1- and 5-min Apgar scores) model was evaluated.

Empirical Study in Twins: Data Collection

Data on Apgar scores and pre- and perinatal characteristics were obtained from the Netherlands Twin Register (NTR)[27]. The NTR recruits families with twins a few weeks to months after birth. Informed consent is obtained from parents. Surveys, including questions on pregnancy, birth and outcomes, were sent to mothers after registration of new-born twins.

Zygosity

For the majority of twin pairs genotyping for zygosity was based on a genome-wide SNP array[30], or on genome-wide sets of micro-satellites. Zygosity typing in earlier studies was based on smaller numbers micro-satellite markers, blood groups[31] or SNPs[27]. For 27% of the same-sex pairs, zygosity was based on items about physical similarity and frequency of confusion of the twins by parents and strangers. These surveys correctly determined the twins' zygosity in 93% of cases[32]. In 19% of the cases zygosity was based on a single item, indicating how much the children look alike, at age 2, which gives a correct determination of zygosity in 92% of the cases[33]. For the other same-sex pairs, zygosity was based on a single question from survey 1.

The sample comprised of 5181 twin pairs, born between 2005 and 2017. Of these, 1763 were MZ and 3418 were DZ (34% and 66%, respectively, reflecting population prevalence in the Netherlands). The dataset includes complete information on zygosity, gestational age, and time between birth of the first and second twin. One-minute Apgar scores were available in 4947 pairs, and 5-min Apgar scores in 4724 pairs. Both Apgar scores were available for 4623 pairs.

The study protocols were approved on March 16, 2004 by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Center, Amsterdam; and May 25, 2017 (NTR-25-mei-2007). All participants provided informed consent.

Variables

Apgar scores were analyzed: (a) as a continuous variable (scores between 0 to 10); (b) as conventional categories (ordinal variables): Apgar values of 0–6 (low), 7–9 (intermediate), and 10 (high); a total score of lower than 7 is considered a source of concern[2].

Based on the literature review, we tested for the effects of birth order (first – second born), zygosity (MZ – DZ), sex (boys – girls), gestational age, birth weight, mother's and father's age at birth, mother's BMI at birth, fetal presentation (head presentation: cephalic; breech and horizontal presentations: non-cephalic), mode of delivery (vaginal and intervention with vacuum extraction, forceps or cesarean section), and intertwin delivery interval. For 1003 MZ twins, we had information on chorionicity[34]. Of these, 745 were monochorionic (MC) and 258 were dichorionic (DC).

Data analyses

Frequencies and means. The data were analysed using SPSS version 25. The frequencies of maternal, delivery and infant characteristics were obtained within each Apgar score

category, for first- and second-born twins, and for MZ and DZ pairs. Differences between continuous variables in MZ and DZ pairs were tested using ANOVA, comparisons between first and second born with paired t-tests.

Fixed effect analysis. The role of maternal, pregnancy, delivery and infant characteristics were analyzed in the first- and second-born twin by ordinal regression. The significant characteristics were included in genetic covariance structure analysis (GCS) of the twin data. Four variables (gestational age, fetal presentation, mode of delivery, birth weight) were selected for inclusion in the analyses of twin resemblance for 1- and 5-min Apgar scores.

Twin correlations. Twin (polychoric) correlations of ordinal Apgar scores in MZ and DZ twins were estimated in Mplus. MZ twins are genetically identical, while DZ twins share on average 50% of their alleles identical by descent (from their parents). The MZ correlation (r_{MZ}) is expected to be greater than the DZ correlations (r_{DZ}), if the phenotype is influenced by genes ($r_{MZ}>r_{DZ}$). The presence of shared environmental factors is suggested if the DZ correlation is larger than half the MZ correlation ($r_{DZ}>r_{MZ}/2$). Unshared environmental influences are present if the MZ correlation is less than one[35].

Genetic covariance structure (GCS) modelling. We carried out GCS analyses of polychoric correlation matrices of the ordinal (3-point) 1- and 5-min Apgar scores using Mplus 6[36]. The analysis of ordinal data is based on the liability-threshold model[37], in which the ordinal scores arise by imposing thresholds on a continuous (standard normal) liability dimension. The twin resemblance at the level of this dimension is expressed by the polychoric correlations. The thresholds are a function of frequencies of the ordinal Apgar values. Given the 3-point ordinal Apgar scores, there are two thresholds in GCS analysis, we first fitted models to estimate the polychoric correlation matrices in the MZ and DZ twins, and to analyse the thresholds, in the presence of the covariates. Subsequently, we fitted an ACE model (see below), in which the phenotypic polychoric correlations are modeled in terms of genetic and shared and unshared environmental effects. Parameter estimates were obtained by means of the weighted least squares estimation (the Mplus estimator WLSMV[36]). Model comparisons were based on the comparison of model Chi-square (goodness of fit) statistics, using the Mplus difference test procedure. The main aim of the GCS modelling was to assess the contributions of genetic and environmental influences to the phenotypic (co-)variance matrix of the 1- and 5-min Apgar scores, while correcting for relevant covariates. We fitted a bivariate ACE model, which included additive genetic (A), shared environmental (C) and unshared environmental effects (E). The results of the analyses provide us with the decomposition of the phenotypic variance of the 1- and 5-min Apgar scores and the decomposition of the phenotype covariance (1-min with 5-min Apgar scores) into genetic and environmental components. In fitting the bivariate ACE model, we used chi-square difference test to test sex differences, and zygosity difference, and birth-order differences in the covariates. We test sex differences in the ACE variance components.

We first fitted a model in which the thresholds differed with respect to birth order (first- vs second-born twin), sex, and zygosity (see **Table S2**). We tested whether the thresholds were equal in males and females (retaining birth order and zygosity-related differences), which was found to be the case, $\chi^2(16)=23.8$, $p=.09$. We tested whether the thresholds

were equal in monozygotic and dizygotic twins, but this was not the case, $\chi^2(8)=23.4$, $p=.003$. So, in the model of choice, we estimated 16 thresholds: 2 for 1-min Apgar scores, 2 for 5-min Apgar scores, which different over zygosity and birth order. In the model, the regression coefficients of the covariates differed with respect to birth order.

We fitted the bivariate ACE model next. In this model we allowed for sex differences in the ACE covariance matrices. As the twin correlation are suggestive of an CE model (i.e., absence of additive genetic effects), we first tested whether we could fix the additive genetic parameters to zero. This was found to be the case: $\chi^2(6)=10.3$, $p=.11$. In this CE model, we constrained the shared by twins and non-shared parameters to be equal over sex, and found sex differences in the shared and non-shared parameters to be absent, $\chi^2(4)=3.958$, $p=.41$.

RESULTS

Review of the literature

The review of studies on characteristics associated with Apgar score included population-based and twin studies published from 1981 to 2018 with exclusion of studies that concern specific clinical groups (e.g., preeclampsia, gestational diabetes), mortality and morbidity in new-borns, and long-term outcomes associated with Apgar scores. Our literature search identified 63 studies, including studies in twins (see **Table 1**) and singletons (see **Table S1** in Supplementary material). The characteristics associated with Apgar scores may be summarized as:

Biological maternal and paternal factors. Maternal short stature[38,39], low maternal age < 17[40], high maternal age >40 year[11,41] and high paternal age >55 years[42] were associated with low Apgar scores. However, Milsom et al.[28] found no association of Apgar scores with maternal age. Higher mother's BMI and maternal obesity were associated with low Apgar scores in some [11,12,43], but not all studies[44,45].

Socioeconomic factors. These were studied in European countries. Maternal occupation, single parent did not show association with Apgar scores in the study of Straube et al.[11]. No significant association was found with employment status during pregnancy[11,28,46]. Low Apgar scores were associated with single motherhood28, missing paternal demographic information[47], low level of mother's education and manual work48,49, and adverse social circumstances[50].

Mode of conception. Two small studies reported an association of 5-min Apgar score with mode of conception[51,52]. No significant difference was found in terms of the 1- and 5-min Apgar scores between twins who were conceived naturally and twins who were conceived with the aid of artificial reproductive technologies[53–58].

Gestational age. Apgar scores were associated with both low[9,20,22,23,39,59–63] and high[12,39] gestational age.

Pregnancy characteristic. A multiple pregnancy is a risk factor of adverse outcomes, including low Apgar scores[12,20,39,58]. This holds specifically for monochorionic

TABLE 1.
Review of studies on prenatal characteristics and Apgar scores in twins

№	Authors	Sample (country)	Sample size n infants (n twin pairs)	Sample and setting	Apgar score (min)	Characteristics associated with low Apgar score (↓)	Characteristics not associated with Apgar score
1	Rayburn et al. (1984)	USA	230 (115 pairs)	Twin sample with vaginal delivery	5 min		Intertwin delivery time
2	Morley et al. (1989)	UK	476 (including 45 twin pairs)	Preterm sample singletons and twins	5 min	Early GA, birth order (second twin)	Multiple birth in preterm
3	Hegyi et al. (1998)	USA	1105 (including 123 twin pairs)	Population-based cohort of preterm births (including multiplies)	1 and 5 min (compo-nents)	Apgar 1 and 5 min: low birthweight, early gestational age, lower arterial blood pH, race (black ↓ vs white). Apgar 1 min: vaginal mode of delivery, sex (male ↓ vs female)	
4	Moise et al. (1998)	Israel	120 (60 pairs)	Case-control: twins after IVF and spontaneous conception	1, 5 min		Mode of conception
5	Daniel et al. (2000)	Israel	594 (297 pairs)	Twin pregnancies conceived spontaneously and via ART	1 min, 5 min	Birth order with mode of conception (↓Apgar 5 min for the first- twin ART-conceived)	
6	Koudstaal et al. (2000)	the Nethe-rlands	288 (144 pairs)	Case-control study: twins after IVF and spontaneous conception	5 min		Mode of conception
7	Thorngren-Jernck et al. (2001)	Sweden	1,028,705 (including 6,433 twin pairs)	Population based cohort, term infants including multiplies	5 min	Vaginal breech delivery, birth weights above 5 kg, second born twins, primiparity, maternal age, smoking, post-date pregnancy, epidural analgesia, male infant gender, being born at night	

Table 1 (continued)

	Authors	Sample (country)	Sample size n infants (n twin pairs)	Sample and setting	Apgar score (min)	Characteristics associated with low Apgar score (↓)	Characteristics not associated with Apgar score
8	Caukwell et al. (2002)	UK	844 (422 pairs)	Twins with different presentation at birth	5 min	Early gestational age	Fetal presentation (cephalic/noncephalic) of the second born with vaginal delivery
9	Usta et al. (2002)	Lebanon	922 (461 pairs)	Twin cohort, term	1 min, 5 min	Birth order (↓ second born)	In second born mode of delivery or presentation
10	Erdemoglu et al. (2003)	Turkey	252 (126 pairs)	Normal twin births	1, 5 min of the second-born	For second twin: early gestational age, low birthweight of the second twin (<1900 g), inter-twin delivery interval >15 min for second twin in breech presentation	Delivery route, fetal presentation
11	Ochsen-kuhn et al. (2003)	Germany	477 (including 78 twin pairs)	Case-control: twins after IVF and spontaneous conception	1, 5, 10 min		Mode of conception
12	Tan et al. (2004)	USA	304,466 (152,233 pairs)	Population-based cohort twins	5 min	Paternal demographic information missing	
13	Wen et al. (2004)	USA	128,219 second twins (128,219 pairs)	Live born second twins	5 min	Birth order and mode of delivery (↓ in second twin delivered with cesarean section after vaginal deli-very of the first twin)	
14	Haest et al. (2005)	the Nether-lands	328 (164 pairs)	Term twin births with different mode of delivery	5 min	Birth order (↓ in second twin)	Mode of delivery (vaginal vs planned cesarean)

Table 1 (continued)

	Authors	Sample (country)	Sample size n infants (n twin pairs)	Sample and setting	Apgar score (min)	Characteristics associated with low Apgar score (↓)	Characteristics not associated with Apgar score
15	Hartley et al. (2005)	USA	10,276 (5,138 pairs)	Twin births with different delivery interval (low risk group)	5 min	Early gestational age; birth order (↓ in second twin); for second twin long intertwin delivery interval >15 min, fetal breech presentation	
16	Usta et al. (2005)	Lebanon	434 (217 pairs)	Twin cohort	1 min, 5 min	Vaginal delivery of vertex-nonvertex twins (↓ in second twin)	
17	Sibony et al. (2006)	France	1,228 (614 pairs)	Twin cohort preterm and term	5 min	Mode of delivery: cesarean section for second born	
18	Bjelic-Radisic et al. (2007)	Austria	562 (281 pairs)	Twin birth cohort	1, 5, 10 min	Birth order (↓ in second twin), for second twin mode of delivery (↓ cesarean section after vaginal delivery of first born or vaginal after vaginal of first born)	Fetal presentation of the second born
19	Sentilhes et al. (2007)	France	412 (206 pairs)	Twin cohort with first twin in breech position, term deliveries	5 min		Mode of delivery (vaginal vs planned cesarean)
20	Herbst et al. (2008)	Sweden	31,982 (15,991 pairs)	Uncomplicated twin pregnancies with different presentation and mode of delivery	5 min	Birth order and mode of delivery (↓ in second born with vaginal delivery)	
21	Schmitz et al. (2008)	France	1,516 (758 pairs)	Twin cohort, term vaginal deliveries, cephalic-presenting first twin	5 min	Mode of delivery: planned vaginal in both twins	

Table 1 (continued)

	Authors	Sample (country)	Sample size n infants (n twin pairs)	Sample and setting	Apgar score (min)	Characteristics associated with low Apgar score (↓)	Characteristics not associated with Apgar score
22	Stein et al. (2008)	Germany	8,220 (4110 pairs)	Population-based twin cohort, term vaginal deliveries of first twin	1, 5, 10 min	Apgar score 1, 5, 10 min for second twin: increased twin-to twin delivery time interval. Apgar score 5 min for second twin: birth weight discordance (birth weight greater in second twin), mode of delivery	Fetal presentation of the 2d born
23	Fox et al. (2010)	USA	574 (287 pairs)	Twin birth with different mode of delivery	1 min, 5 min	Non-active second-stage management	Mode of delivery (vaginal vs planned cesarean)
24	Kwon et al. (2011)	Korea	158 (79 pairs)	Twin births with known umbilical arterial blood parameters	1 min, 5 min	Vaginal delivery and intertwin delivery time (↓ for second twin)	
25	Schneuber et al. (2011)	Austria	414 (207 pairs)	Twin births, term deliveries	1, 5, 10 min		Intertwin delivery time
26	Barrett et al. (2013)	Canada	5,607 (2,804 pairs)	Twin births	5 min		Mode of delivery (planned vaginal vs planned cesarean section) on low Apgar as the component of primary outcome
27	Fan et al. (2013)	China	750 (375 pairs)	Dichorionic twin pregnancies conceived spontaneously and via ART	1 min, 5 min		Mode of conception (spontaneous vs ART)
28	Caserta et al. (2014)	Italy	690 (345 pairs)	Dichorionic, diamniotic twin pregnancies conceived spontaneously and via ART	5 min		Mode of conception (spontaneous vs ART)

Table 1 (continued)

	Authors	Sample (country)	Sample size n infants (n twin pairs)	Sample and setting	Apgar score (min)	Characteristics associated with low Apgar score (↓)	Characteristics not associated with Apgar score
29	Hjorto et al. (2014)	Denmark	1,100 (550 pairs)	Twin deliveries	1 min	For second born chorionicity (monochorionic), time between birth, vacuum extraction	
30	Vogel J.P. et al. (2014)	WHO Global Survey on Maternal Health (multi-country Africa, Asia, Latin America)	2,848 (1,424 pairs)	Multiple births	5 min	For second twin: non-vertex presentation after vaginal delivery of first twin	
31	Wenckus et al. (2014)	USA	4,450 (2,225 pairs)	Twin birth cohort	5 min	Mode of delivery (vaginal delivery ↓ vs cesarian section) in twins	
32	Dolgun et al. (2016)	Turkey	176 (88 pairs)	Preterm twins without complications	1 min, 5 min	Apgar score 1 min: low fetal body weight, early gestational age, low height, head circumference	Apgar 1 and 5 min: maternal age, placental weight, length of umbilical cord, premature rupture of membranes, birth order. Apgar 5 min: gender, mode of delivery, gestational age, height, head circumference.
33	Jhaveri et al. (2016)	India	186 (93 pairs)	Twin birth sample	5 min	For second born vaginal mode of delivery	

Table 1 (continued)

	Authors	Sample (country)	Sample size n infants (n twin pairs)	Sample and setting	Apgar score (min)	Characteristics associated with low Apgar score (↓)	Characteristics not associated with Apgar score
34	Machado et al. (2017)	Portugal	1051 (540 pairs)	Twin preg-nancies with known chorionicity	5 min		Chorionicity
35	Vulic et al. (2017)	Croatia	434 (217 pairs)	Dichorionic twin pregnancies	Not indicated	Discordant twin growth in DC pregnancies	
36	Pourali et al. (2016)	Iran	254 (127 pairs)	Dichorionic twin pregnancies following ART vs. spontaneous	1 min, 5 min	Multiple pregnancy	Mode of conception (ART vs spontaneous)
37	Zhao et al. (2017)	USA	216,076 (108,038 pairs)	Mixed-gender twin pairs	5 min	Male sex, mixed-gender twin pairs	
38	Algeri et al. (2018)	Italy	800 (400 pairs)	Diamniotic pregnancies	Not indicated		Intertwin delivery time
39	Lindroos et al. (2018)	Sweden	1054 (527 pairs)	Twin birth with different delivery interval (not high risk group)	5 min	Lower birthweight, early gestational age, inter-twin birthweight discordance >25%, chorionicity (MC ↓ for second born)	For second twin twin-to-twin time interval (30 min or >30 min), presentation at birth

Note: studies are indicated with ** in the reference list. MC, monochorionic; DC, dichorionic.

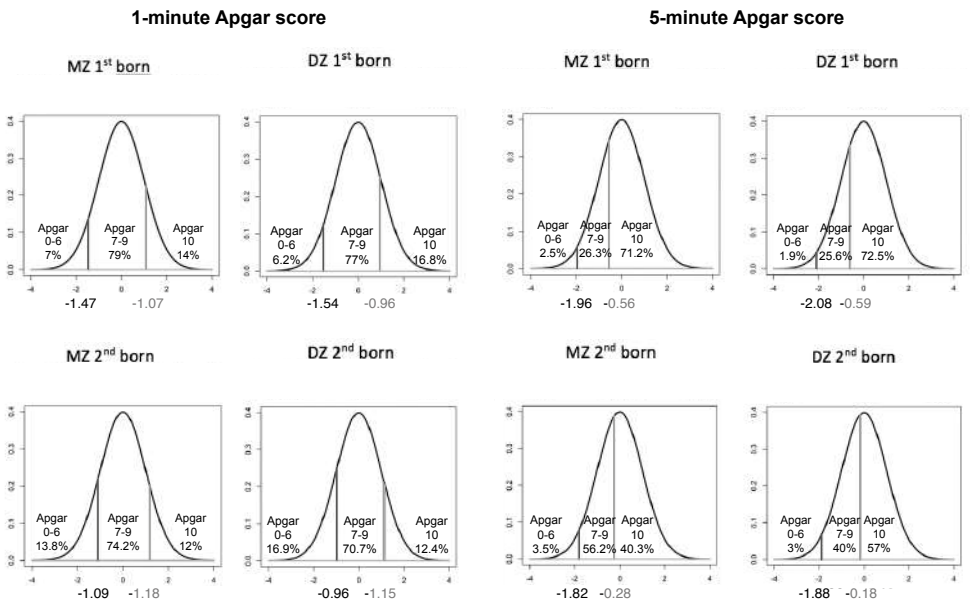


Figure 1. Proportions of MZ and DZ newborn twins in different categories of 1- and 5-min Apgar score

pregnancies[62,64] and dichorionic pregnancies with discordant fetal weight[65]. Again, some studies failed to find an association with chorionicity[66]. Even though maternal smoking is generally associated with negative outcomes in neonates, most studies found no significant association between Apgar scores and prenatal maternal smoking after accounting for other confounders[11,28,50,67], although an association between mother's smoking during the first trimester and low Apgar score was seen by Kallen[68].

Delivery characteristics. Multiple deliveries were associated with adverse outcomes in the second-born twin[12,18–21,69,70]. The intertwin delivery interval is an important determinant of the adverse effects on the second-born[60,61,64,69,71] as this interval is related to the risk of hypoxia, due to decreasing pH in the umbilical arterial blood. However, other studies showed that even a relatively long intertwin delivery interval was not associated with unfavourable Apgar scores[28,62,72–74].

Non-cephalic (breech and horizontal) presentation at birth is associated with low Apgar score in singletons[75–77] and twins[61]. After a vaginal delivery of a vertex first twin, non-vertex presentation of the second twin was associated with increased odds of a low 5-min Apgar score[77]. However, other studies found no support for an effect of fetal presentation on the Apgar score of the second twin[59,62,70,78,79]. Emergency interventions (vacuum extraction, forceps, urgent operative delivery) were associated with adverse outcomes, including low Apgar scores[28,45,64]. The risk of a low Apgar score given at planned vaginal delivery was much higher than the risk associated with a selective caesarean section in singletons[22,75] and twins[79], especially in second-born twin[19,69,80–82]. The effect of mode of delivery in twins was not supported by some studies analyzing different fetal presentation

deliveries[83–85]. Cesarean section was associated with low Apgar in second twin following vaginal delivery of first twin[21,29].

New-born characteristics. Several studies found a positive association between birth weight and Apgar scores[12,22,23,60,62,71,86]. However, Iliodromiti et al[9] found no association of birth weight and Apgar scores in the large population-based sample of more than 1 million births in Scotland. Birthweight discordance in twins is associated with low Apgar scores for the second-born[62,71]. On average, girls have higher Apgar scores than boys[22,23,87,88].

Based on our literature review, we included the following risk factors in our analyses of the twin data: zygosity, chorionicity, birth order, gestational age, birth weight, sex, mother's age and father's age at birth, mother's BMI at birth, mode of delivery, and fetal presentation.

Descriptives and Apgar scores in twins

Proportions of new-born NTR twins in three categories of 1- and 5-min Apgar scores (low, intermediate, and high) for first and second-born in MZ and DZ twin pairs are presented in **Figure 1** (distribution in pairs see in **Table S3**). In terms of the proportions, the first- and second-born twins do not differ greatly with respect to the 1-min Apgar scores. For instance, the proportion of intermediate 1-minute Apgar score (7–9) are about 79%, 74%, 77%, and 71% (MZ first- born, MZ second born, DZ first-born, and DZ second-born, respectively). There are however appreciable differences between the first and second-born twins in the 5-min Apgar scores. For instance, the proportion of a high 1-min Apgar score (10) are about 71%, 40%, 72%, and 57% (MZ first-born, MZ second-born, DZ first-born, DZ second-born, respectively).

We compared the means of Apgar scores in MZ and DZ twin pairs between first- and second-born twins. Mean continuous Apgar scores were higher in the first-born than in second-born twins both in MZ and DZ pairs: in MZ at first minute 8.45 vs. 8.1 (p<.0001), at fifth minute 9.36 vs. 9.22 (p<.0001); in DZ at first minute 8.61 vs. 8.03 (p<.0001), at fifth minute 9.51 vs. 9.26 (p<.0001). MZ first- born twins had lower Apgar scores than DZ first-born twins: at first minute 8.43 vs 8.6 in DZ, at fifth minute 9.36 vs 9.51 in DZ (p<.0001). In second-born twins, there was no effect of zygosity (see Table S4).

Next, we tested pre- and perinatal characteristics for the first- and second-born twins. The distribution of perinatal and delivery characteristics of NTR twins are presented in **Table 2** (for information on characteristics of low, intermediate and high Apgar score groups see **Table S5** in the Supplementary material). Given alpha of .05, mother's age, father's age, and mother's BMI at birth did not predict Apgar scores (**Table S6** in the Supplementary materials). The intertwin delivery time was significant for Apgar scores of the second-born twin. There was no significant effect of chorionicity on Apgar score in MZ twins with chorionicity data, taking into account gestational age, birth weight, sex, mode of delivery, and fetal presentation (**Table S7** in the Supplementary materials). Characteristics without significant effects for both twins were excluded from further analysis.

Multi-group genetic covariance structural equation modelling of the ordinal 1- and 5-min Apgar scores included monozygotic male (MZM), and female (MZF), dizygotic male (DZM), and female (DZF), dizygotic male-female (DZMF) and dizygotic female-

TABLE 2.
Prenatal, delivery, and infant characteristics of monozygotic and dizygotic twin pairs

	MZ twin pairs (*n*=1763)		DZ twin pairs (3418)	
	first-born	**second-born**	**first-born**	**second-born**
Sex				
male	879 (49,9%)	879 (49,90%)	1724 (50,40%)	1723 (50,40%)
female	884 (50,10%)	884 (50,10%)	1694 (49,60%)	1695 (49,60%)
Total	1763	1763	3418	3418
Gestational age (weeks)				
<=32	126 (7,10%)		148 (4,30%)	
33-36	916 (52,00%)		1294 (37,90%)	
>=37	721 (40,90%)		1976 (57,80%)	
Total	1763		3418	
Mother's age at birth (years)				
18-29	651 (36,90%)		877 (25,70%)	
30-39	1056 (59,90%)		2426 (71,00%)	
>40	56 (3,20%)		115 (3,40%)	
Total	1763		3418	
Mother's BMI at birth				
<25	177 (10,80%)		293 (9,20%)	
25-30	726 (44,40%)		1346 (42,30%)	
>30	733 (44,80%)		1544 (48,50%)	
Total	1636		3183	
missing	127		235	
Father's age at birth (years)				
20-29	353 (20,60%)		461 (14,00%)	
30-39	1164 (67,80%)		2300 (69,80%)	
>40	199 (11,60%)		532 (16,20%)	
Total	1716		3293	
missing	47		125	
Mode of conception				
naturally	1592 (91,10%)		1939 (57,30%)	
stimulated	55 (3,10%)		531 (15,70%)	
IVF/ICSI	101 (5,80%)		913 (27,00%)	
Total	1748		3383	
missing	15		35	
Birthweight				
<1500	102 (5,80%)	112 (6,40%)	130 (3,80%)	158 (4,70%)
1500-2500	821 (46,70%)	869 (49,50%)	1245 (36,50%)	1352 (39,80%)
>2500	835 (47,50%)	775 (44,10%)	2033 (59,70%)	1885 (55,50%)
Total	1758	1756	3408	3395
missing	5	7	10	23
Fetal presentation				
cephalic	1487 (84,70%)	1032 (59,40%)	2562 (75,30%)	1758 (52,10%)
noncephalic (breech, horizontal)	269 (15,30%)	706 (40,60%)	842 (24,70%)	1617 (47,90%)
Total	1756	1738	3404	3375
missing	7	25	14	43

Table 2 (continued)

	MZ twin pairs (*n*=1763)		DZ twin pairs (3418)	
	first-born	**second-born**	**first-born**	**second-born**
Mode of delivery				
vaginal	970 (55,20%)	968 (55,50%)	1734 (50,90%)	1617 (47,80%)
planned caesarean section	235 (13,40%)	234 (13,40%)	650 (19,10%)	649 (19,20%)
urgent intervention (forceps, vacuum extraction)	182 (10,40%)	110 (6,30%)	314 (9,20%)	257 (7,60%)
urgent caesarean section	370 (21,10%)	431 (24,70%)	712 (20,90%)	863 (25,50%)
Total	1757	1743	3410	3386
missing	6	20	8	32
Intertwin delivery time (minutes)				
<5	767 (43,50%)		1492 (43,70%)	
6-15	612 (34,70%)		914 (26,70%)	
16-60	342 (19,40%)		866 (25,30%)	
61-1440	42 (2,40%)		146 (4,30%)	
Total	1763		3418	

Note: values are n (%).

male (DZFM) pairs. We included gestational age (a characteristic of twin pairs) and fetal presentation, birth weight, mode of delivery (characteristics of individual twins) as covariates. We estimated the effects of covariates and the polychoric twin correlations (**Table S2** in the Supplementary material).

In multi-group genetic analyses the effects of covariates did not differ with respect to sex (χ^2 (16)=8.88, p=.91). Gestational age had a positive effect on both Apgar measurements in both twins (p<.0001) (**Table 3**). Birth weight had a positive effect on 5-min Apgar score of the first born ($ß_{1min}$=.09, p=.002). The effects of delivery characteristics, such as mode of delivery and fetal presentation at birth, were different for first- and second-born twins. Non-cephalic presentation at birth of the first-born twin had a positive effect on Apgar scores of the first-born ($ß_{1min}$=.11, p=.02; $ß_{5min}$=.19, p<.0001), and non-cephalic presentation of the second-born twin have negative effect on Apgar scores ($ß_{1min}$=-.23, p<.0001; $ß_{5min}$=-.16, p<.0001). First-born twins delivered vaginally were more likely to have higher Apgar scores at both points ($ß_{1min}$=.26, $ß_{5min}$=.42, p<.0001). Second-born twins delivered vaginally were more likely to have lower 1-min Apgar scores ($ß_{1min}$=-.14, p<.0001) and the effect was not significant for the 5-min Apgar score ($ß_{5min}$=-0.002, ns).

Table 4 summarizes the twin correlations with and without the correction for covariates (see also **Table S8** in the Supplemental material). Overall, the correlations, which varied between .43 and .55, did not differ greatly between zygosities, which suggests the absence of genetic effects. Shared environmental influences accounted for 52.6% (CI-95% [0.51, 0.54]) and 50.2% (CI-95% [0.48, 0.52]) of the variance of the 1- and 5-min Apgar scores, respectively. The remainder of the variance was explained by non-shared environmental effects: 47.4% (CI 95% [0.46, 0.49]) and 49.8% (CI 95% [0.48,

TABLE 3.
Average for continuous Apgar scores and effect of gestational age, birth weight, fetal presentation and mode of delivery on ordinal 1- and 5-min Apgar score in first and second-born twins

	ß	SE	Est./SE	*p*	CI (95%)
1-min Apgar first-born twin: mean 8.45					
Effect size for ordinal score					
Gestational age (Z-scores)	0.245	0.032	7.483	0.000	[0.161; 0.329]
Birthweight (Z-scores)	0.043	0.032	1.431	0.153	[-0.038; 0.125]
Fetal presentation (0=cephalic/1=non-cephalic)	0.110	0.048	2.322	0.020	[-0.014; 0.233]
Mode of delivery (0=vaginal/1=intervention)	-0.258	0.041	-6.297	0.000	[-0.363; -0.152]
5-min Apgar first-born twin: mean 9.36					
Effect size for ordinal score					
Gestational age (Z-scores)	0.301	0.033	9.000	0.000	[0.215; 0.387]
Birthweight (Z-scores)	0.097	0.032	3.055	0.002	[0.014; 0.179]
Fetal presentation (cephalic/non-cephalic)	0.194	0.051	3.818-	0.000	[0.062; 0.325]
Mode of delivery (vaginal/intervention)	-0.417	0.041	10.263	0.000	[-0.522; -0.312]
1-min Apgar second-born twin: mean 8.61					
Effect size for ordinal score					
Gestational age (Z-scores)	0.279	0.031	8.991	0.000	[0.200; 0.360]
Birthweight (Z-scores)	-0.001	0.030	-0.012-	0.990	[-0.078; 0.075]
Fetal presentation (cephalic/non-cephalic)	-0.231	0.033	7.042	0.000	[-0.315; -0.147]
Mode of delivery (vaginal/intervention)	0.141	0.034	4.100	0.000	[0.052; 0.229]
5-min Apgar second-born twin: mean 9.51					
Effect size for ordinal score					
Gestational age (Z-scores)	0.346	0.032	10.948	0.000	[0.265; 0.428]
Birthweight (Z-scores)	0.026-	0.030	0.894-	0.372	[-0.052; 0.104]
Fetal presentation (cephalic/non-cephalic)	0.163	0.033	4.966	0.000	[-0.248; -0.079]
Mode of delivery (vaginal/intervention)	0.002	0.034	0.070	0.944	[-0.086; 0.091]

Note: SE=standard error, Est.=estimates, and CI=confidential interval

0.52]) of the variance of the 1- and 5-min Apgar scores, respectively. The correlation between the 1- and 5-min Apgar scores was .70 (CI 95% [0.68, 0.72]), which consistent with the correlations shown in **Table 4**. This correlation is decomposed into .33 (CI 95% [0.30, 0.35]) due to shared prenatal environmental factors, and .37 (CI-95% [0.39, 0.40]) due to unshared prenatal environmental factors (see **Table S9** and **Figure S1** in the Supplemental material).

DISCUSSION

Our literature review identified characteristics that were significantly associated with Apgar scores in population-based and twin studies. These included gestational age, birth weight, sex, mother's and father's age at birth, mother's BMI, mode of delivery, and fetal presentation, specifically for twins zygosity, chorionicity, birth order, and intertwin delivery time. In the current analyses of twins, birth order, zygosity, gestational age, birth weight, fetal presentation at birth, and mode of delivery contributed to Apgar scores.

TABLE 4.
Twin correlations for 1- and 5-min Apgar score (ordinal variables) noncorrected and corrected for gestational age, birth weight, mode of delivery and fetal presentation

		r	*r* adjusted to covariates
Twin correlations 1-min Apgar			
	MZM	.621	.552
	MZF	.607	.537
	DZM	.544	.485
	DZF	.610	.551
	DZMF	.539	.479
	DZFM	.595	.536
Twin correlations 5-min Apgar			
	MZM	.676	.540
	MZF	.643	.507
	DZM	.630	.518
	DZF	.620	.508
	DZMF	.568	.455
	DZFM	.547	.435

Note: MZM = monozygotic male, MZF = monozygotic female, DZM = dizygotic male, DZF = dizygotic female, DZMF = dizygotic male–female and DZFM = dizygotic female–male.

In our empirical study, parental characteristics were not associated with Apgar scores. The effects of mother's and father's age on Apgar scores as established in other studies may be explained by families included in these other studies, such as teenage mothers[40], mothers over age >40[41], and older fathers[42]. In our sample, only 171 women were above 40 years (3.2% MZ and 3.4% DZ of mothers), and there were no mothers younger than 18. An effect of a maternal BMI on Apgar score was found in some studies of singletons[11,12,43]. Multiple pregnancies are generally accompanied by greater maternal BMI due to gestational weight gain than in singleton pregnancies. This may explain the non-significance of maternal BMI effect on Apgar score in our study (near 45% of mothers in our sample had BMI >30). Optimal gestational weight gain in twin pregnancy is unclear[89]. However, some studies also support our findings that maternal age[28] and BMI at birth[44,45] are not associated with Apgar score. Monochorionicity was an important risk factor for adverse perinatal outcomes in twins[34,62,64,90], but was not associated with Apgar scores in our study.

In accordance with many singleton[9,22,39,63] and twin[20,23,59–62] studies, we found a large effect of gestational age. In premature new-borns, a low Apgar score may indicate intrinsic physiological immaturity and inadequate capacity for response rather than abnormal physiological functions[9]. Preterm twins have the same prognosis as preterm singletons[20]. The effect of gestational age was stronger than birth weight. We found effects of birth weight on Apgar scores in first- born twins, but not in the second-born.

Our findings on birth order agree with previous findings. The first twin is in better clinical condition[12,18–21,25,69,70]. The second twin is at greater risk of lower scores, which can be due to longer delivery time, risk of hypoxia, non-definable fetal presentation before birth

to decide the better tactics, or complications during delivery. The proportion of low Apgar score in first- and second-born reported by Franchi-Pinto et al.[25] corresponds with our findings.

Fetal presentation at birth showed opposite effects in first- and second-born twins and should be further investigated together with mode of delivery and in application of horizontal/non-horizontal classification of presentation. The positive effect of non-cephalic presentation in first-born twins in our study is in contrast with other studies[61], and could be associated with tactics of delivery that can be planned in comparison with delivery of the second-born. Delivery practice in case of non-cephalic presentation of the first-born can increase the probability of high Apgar scores in the new-borns. The previous studies have shown that the effect of fetal presentation on Apgar scores in second-born is associated with fetal presentation and mode of delivery of the first-born[59,62,70,77–79]. Cephalic presentation of the second twin is associated with higher Apgar scores in our study, in line with previous studies.

Physicians have gained a clear understanding of how to deliver twins with regard to their presentation and gestation. Some retrospective analyses and meta-analyses reported that the prognosis of twins was not different according to delivery mode[29,70,91], but population-based studies reported that the mortality rate or complications in second twins were higher in vaginal deliveries[19,21,24,69,79–82,92]. We found better outcomes, in terms of Apgar score, for vaginal delivery in first-born twins and intervention delivery in second-born. We did not confirm that caesarean section is associated with low Apgar in second-born twins as shown by Wen et al[21]. The mode of twin delivery should be considered on the basis of information on fetal presentation of both twins; when concerning the mode of delivery of the second-born one should take in account the mode of delivery of the first-born.

To our knowledge, ours was the first large study evaluating genetic influences on 1- and 5-min Apgar scores. In contrast to a smaller twin study done without precise zygosity definition[25], we did not find evidence for genetic influences on Apgar scores. The slightly higher correlations in monozygotic twins in our study partly corresponds with the intraclass correlation coefficients reported by Riese[26] for 1-min Apgar scores in a small sample of monozygotic and dizygotic twins. We did observe large influences of non-genetic factors shared by twins from the same pairs. We acknowledge that a shared environmental component could reflect to some extent a shared measurement bias (e.g. if both twins are rated at the same time by one nurse). Apgar scores represent routine clinical practice, but some of the variability could reflect heterogeneity in clinical scoring practices as opposed to true differences in biomedical outcomes[14]. Also, the genotype of the mother in part creates the prenatal environment of both twins and thus is part of the ‘shared environment’.

Our data do not reflect the whole population as it does not include cases with infant death. If the individual components of Apgar score (skin color or appearance, pulse rate, reflex, activity and respiratory effort) would be available for analyses, it is possible that the contribution of shared and non-shared environment and genetic influence would differ across components. Twin-specific in-utero environment and epigenetic factors

also are of interest for future studies to examine the sources of unique environment. For understanding the variance of shared and non-shared perinatal environment further analysis of mother's health status and early medical support is needed.

CONCLUSIONS

We have found that for both MZ and DZ pairs second-born twins have lower Apgar scores in comparison with first-born twins. There are different effects of pre- and perinatal characteristics on 1- and 5-min Apgar score in first- and second-born twins. Based on twin analyses a genetic component was not significant for Apgar scores. For 1- and 5-min Apgar score, about half of the variation was explained by shared and half by non-shared environmental factors. It is possible that some of the shared environment is due to the same rater scoring both twins. The most important factors for Apgar scores are gestational age, birth weight, birth order, zygosity, fetal presentation, and mode of delivery.

References

*with ** are indicated studies on Apgar scores in twins (presented in Table 1) and with * are indicated studies on Apgar scores in singletons (presented in Table S1 in the Supplemental material)*

1. Apgar, V. A proposal for a new method of evaluation of the newborn infant. *Current researches in anesthesia & analgesia* **32**, 260–267 (1953).
2. Committee Opinion No. 644: The Apgar Score. *Obstetrics and gynecology* **126**, e52–e55 (2015).
3. Apgar, V. The newborn (Apgar) scoring system. Reflections and advice. *Pediatric clinics of North America* **13**, 645–650 (1966).
4. Drage, J. S., Kennedy, C. & Schwarz, B. K. The Apgar Score as an index of neonatal mortality. A report from the Collaborative Study of Cerebral Palsy. *Obstetrics and gynecology* **24**, 222–230 (1964).
5. Harrington, D. J., Redman, C. W., Moulden, M. & Greenwood, C. E. The long-term outcome in surviving infants with Apgar zero at 10 minutes: a systematic review of the literature and hospital-based cohort. *American journal of obstetrics and gynecology* **196**, 463.e1–5 (2007).
6. Nelson, K. B. & Ellenberg, J. H. Apgar scores as predictors of chronic neurologic disability. *Pediatrics* **68**, 36–44 (1981).
7. Casey, B. M., McIntire, D. D. & Leveno, K. J. The continuing value of the Apgar score for the assessment of newborn infants. *The New England journal of medicine* **344**, 467–471 (2001).
8. Grünebaum, A. *et al.* Apgar score of 0 at 5 minutes and neonatal seizures or serious neurologic dysfunction in relation to birth setting. *American journal of obstetrics and gynecology* **209**, 323.e1–6 (2013). *
9. Iliodromiti, S., Mackay, D. F., Smith, G. C. S., Pell, J. P. & Nelson, S. M. Apgar score and the risk of cause-specific infant mortality: a population-based cohort study. *Lancet (London, England)* **384**, 1749–1755 (2014). *
10. Li, F. *et al.* The apgar score and infant mortality. *PloS one* **8**, e69072 (2013).
11. Straube, S. *et al.* Investigation of the association of Apgar score with maternal socio-economic and biological factors: an analysis of German perinatal statistics. *Archives of gynecology and obstetrics* **282**, 135–141 (2010). *
12. Thorngren-Jerneck, K. & Herbst, A. Low 5-minute Apgar score: a population-based register study of 1 million term births. *Obstetrics and gynecology* **98**, 65–70 (2001).**
13. *Guidelines on Basic Newborn Resuscitation. Guidelines on Basic Newborn Resuscitation* (World Health Organization, 2012).
14. Siddiqui, A. *et al.* Can the Apgar Score be Used for International Comparisons of Newborn Health? *Paediatric and perinatal epidemiology* **31**, 338–345 (2017).
15. O'Donnell, C. P. F., Kamlin, C. O. F., Davis, P. G., Carlin, J. B. & Morley, C. J. Interobserver variability of the 5-minute Apgar score. *The Journal of pediatrics* **149**, 486–489 (2006).
16. Rüdiger, M. & Konstantelos, D. Apgar score and risk of cause-specific infant mortality. *Lancet (London, England)* vol. 385 505–506 (2015).
17. Kattwinkel, J. *et al.* Part 15: neonatal resuscitation: 2010 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. *Circulation* **122**, S909–19 (2010).
18. Haest, K. M. J., Roumen, F. J. M. E. & Nijhuis, J. G. Neonatal and maternal outcomes in twin gestations ≥32 weeks according to the planned mode of delivery. *European Journal of Obstetrics & Gynecology and Reproductive Biology* **123**, 17–21 (2005).**
19. Herbst, A. & Källén, K. Influence of mode of delivery on neonatal mortality in the second twin, at and before term. *BJOG : an international journal of obstetrics and gynaecology* **115**, 1512–1517 (2008).**
20. Morley, R., Cole, T. J., Powell, R. & Lucas, A. Growth and development in premature twins. *Archives of disease in childhood* **64**, 1042–1045 (1989).**

21. Wen, S. W. *et al.* Neonatal morbidity in second twin according to gestational age at birth and mode of delivery. *American journal of obstetrics and gynecology* **191**, 773–777 (2004).**

22. Hegyi, T. *et al.* The apgar score and its components in the preterm infant. *Pediatrics* **101**, 77–81 (1998).**

23. Dolgun, Z. N., Inan, C., Altintas, A. S., Okten, S. B. & Sayin, N. C. Preterm birth in twin pregnancies: Clinical outcomes and predictive parameters. *Pakistan journal of medical sciences* **32**, 922–926 (2016).**

24. Rossi, A., Mullin, P. & Chmait, R. Neonatal outcomes of twins according to birth order, presentation and mode of delivery: a systematic review and meta-analysis*. *BJOG: An International Journal of Obstetrics & Gynaecology* **118**, 523–532 (2011).

25. Franchi-Pinto, C., Colletto, G. M. D. D., Krieger, H. & Beiguelman, B. Genetic effect on Apgar Score. *Genetics and Molecular Biology* **22**, 13–16 (1999).

26. Riese, M. L. Genetic influences on neonatal temperament. *Acta geneticae medicae et gemellologiae* **39**, 207–213 (1990).

27. van Beijsterveldt, C. E. M. *et al.* The Young Netherlands Twin Register (YNTR): Longitudinal Twin and Family Studies in Over 70,000 Children. *Twin Research and Human Genetics* **16**, 252–267 (2013).

28. Milsom, I. *et al.* Influence of maternal, obstetric and fetal risk factors on the prevalence of birth asphyxia at term in a Swedish urban population. *Acta obstetricia et gynecologica Scandinavica* **81**, 909–917 (2002). *

29. Sibony, O., Touitou, S., Luton, D., Oury, J.-F. & Blot, P. Modes of delivery of first and second twins as a function of their presentation. Study of 614 consecutive patients from 1992 to 2000. *European journal of obstetrics, gynecology, and reproductive biology* **126**, 180–185 (2006).**

30. Odintsova, V. v. *et al.* Establishing a Twin Register: An Invaluable Resource for (Behavior) Genetic, Epidemiological, Biomarker, and 'Omics' Studies. *Twin Research and Human Genetics* **21**, 239–252 (2018).

31. Van Dijk, B. A., Boomsma, D. I. & de Man, A. J. Blood group chimerism in human multiple births is not rare. *American journal of medical genetics* **61**, 264–268 (1996).

32. Rietveld, M. J. *et al.* Zygosity diagnosis in young twins by parental report. *Twin research: the official journal of the International Society for Twin Studies* **3**, 134–141 (2000).

33. Groen-Blokhuis, M. M., Middeldorp, C. M., van Beijsterveldt, C. E. M. & Boomsma, D. I. Evidence for a causal association of low birth weight and attention problems. *Journal of the American Academy of Child and Adolescent Psychiatry* **50**, 1247–54.e2 (2011).

34. Van Beijsterveldt, C. E. M. *et al.* Chorionicity and Heritability Estimates from Twin Studies: The Prenatal Environment of Twins and Their Resemblance Across a Large Number of Traits. *Behavior genetics* **46**, 304–314 (2016).

35. Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. *Nature reviews. Genetics* **3**, 872–882 (2002).

36. Muthén, L. K. & Muthén, B. O. *Mplus User's Guide.* (Muthén & Muthén, 2007).

37. Falconer, D. Quantitative genetics in Edinburgh: 1947–1980. *Genetics* **133**, 137–142 (1993).

38. Camilleri, A. P. The obstetric significance of short stature. *European journal of obstetrics, gynecology, and reproductive biology* **12**, 347–356 (1981). *

39. Svenvik, M., Brudin, L. & Blomberg, M. Preterm Birth: A Prominent Risk Factor for Low Apgar Scores. *BioMed Research International* **2015**, 1–8 (2015). *

40. Chen, X.-K. *et al.* Teenage pregnancy and adverse birth outcomes: a large population based retrospective cohort study. *International journal of epidemiology* **36**, 368–373 (2007). *

41. Jahromi, B. N. & Husseini, Z. Pregnancy outcome at maternal age 40 and older. *Taiwanese journal of obstetrics & gynecology* **47**, 318–321 (2008). *

42. Khandwala, Y. S. *et al.* Association of paternal age with perinatal outcomes between 2007 and 2016 in the United States: population based cohort study. *BMJ (Clinical research ed.)* **363**, k4372 (2018). *

43. Chen, M. *et al.* Maternal obesity and neonatal Apgar scores. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians* **23**, 89–95 (2010). *

44. Usha Kiran, T. S., Hemmadi, S., Bethel, J. & Evans, J. Outcome of pregnancy in a woman with an increased body mass index. *BJOG : an international journal of obstetrics and gynaecology* **112**, 768–772 (2005). *

45. Rode, L., Nilas, L., Wøjdemann, K. & Tabor, A. Obesity-related complications in Danish single cephalic term pregnancies. *Obstetrics and gynecology* **105**, 537–542 (2005). *

46. Marbury, M. C. *et al.* Work and pregnancy. *Journal of occupational medicine. : official publication of the Industrial Medical Association* **26**, 415–421 (1984). *

47. Tan, H., Wen, S. W., Walker, M. & Demissie, K. Missing paternal demographics: A novel indicator for identifying high risk population of adverse pregnancy outcomes. *BMC pregnancy and childbirth* **4**, 21 (2004).**

48. Hemminki, E., Malin, M. & Rahkonen, O. Mother's social class and perinatal problems in a low-problem area. *International journal of epidemiology* **19**, 983–990 (1990). *

49. Odd, D. E. *et al.* Risk of low Apgar score and socioeconomic position: a study of Swedish male births. *Acta paediatrica (Oslo, Norway: 1992)* **97**, 1275–1280 (2008). *

50. Kalland, M., Sinkkonen, J., Gissler, M., Meriläinen, J. & Siimes, M. A. Maternal smoking behavior, background and neonatal health in Finnish children subsequently placed in foster care. *Child abuse & neglect* **30**, 1037–1047 (2006). *

51. Daniel, Y. *et al.* Analysis of 104 twin pregnancies conceived with assisted reproductive technologies and 193 spontaneously conceived twin pregnancies. *Fertility and sterility* **74**, 683–689 (2000).**

52. Ramoğlu, M. G., Kavuncuoğlu, S., Özbek, S. & Aldemir, E. Perinatal and somatic growth properties of preterm babies born from spontaneous and in vitro fertilization multiple pregnancies. *Turk pediatri arsivi* **49**, 17–24 (2014). *

53. Caserta, D. *et al.* Maternal and perinatal outcomes in spontaneous versus assisted conception twin pregnancies. *European journal of obstetrics, gynecology, and reproductive biology* **174**, 64–69 (2014).**

54. Fan, C., Sun, Y., Yang, J., Ye, J. & Wang, S. Maternal and neonatal outcomes in dichorionic twin pregnancies followingIVF treatment: a hospital-based comparative study. *International journal of clinical and experimental pathology* **6**, 2199–2207 (2013).**

55. Koudstaal, J. *et al.* Obstetric outcome of twin pregnancies after in-vitro fertilization: a matched control study in four Dutch university hospitals. *Human reproduction (Oxford, England)* **15**, 935–940 (2000).**

56. Moise, J., Laor, A., Armon, Y., Gur, I. & Gale, R. The outcome of twin pregnancies after IVF. *Human reproduction (Oxford, England)* **13**, 1702–1705 (1998).**

57. Ochsenkühn, R. *et al.* Pregnancy complications, obstetric risks, and neonatal outcome in singleton and twin pregnancies after GIFT and IVF. *Archives of gynecology and obstetrics* ***268***, 256–261 (2003).**

58. Pourali, L., Ayati, S., Jelodar, S., Zarifian, A. & Sheikh Andalibi, M. S. Obstetrics and perinatal outcomes of dichorionic twin pregnancy following ART compared with spontaneous pregnancy. *International journal of reproductive biomedicine* **14**, 317–322 (2016).**

59. Caukwell, S. & Murphy, D. J. The effect of mode of delivery and gestational age on neonatal outcome of the non-cephalic- presenting second twin. *American journal of obstetrics and gynecology* **187**, 1356–1361 (2002).**

60. Erdemoglu, E., Mungan, T., Tapisiz, O. L., Ustunyurt, E. & Caglar, E. Effect of inter-twin delivery time on Apgar scores of the second twin. *The Australian & New Zealand journal of obstetrics & gynaecology* **43**, 203–206 (2003).**

61. Hartley, R. S. & Hitti, J. Birth order and delivery interval: Analysis of twin pair perinatal outcomes. *The Journal of Maternal-Fetal & Neonatal Medicine* **17**, 375–380 (2005).**

62. Lindroos, L., Elfvin, A., Ladfors, L. & Wennerholm, U.-B. The effect of twin-to-twin delivery time intervals on neonatal outcome for second twins. *BMC pregnancy and childbirth* **18**, 36 (2018).**

63. Van der Ven, A. J. *et al.* Comparison of Perinatal Outcome of Preterm Births Starting in Primary Care versus Secondary Care in Netherlands: A Retrospective Analysis of Nationwide Collected Data. *Obstetrics and gynecology international* **2014**, 423575 (2014).

64. Hjortø, S., Nickelsen, C., Petersen, J. & Secher, N. J. The effect of chorionicity and twin-to-twin delivery time interval on short-term outcome of the second twin. *The journal of maternal-fetal & neonatal medicine: the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians* **27**, 42–47 (2014).**

65. Vulić, M. *et al.* A Retrospective Study of Discordant Twin Growth in Dichorionic Twin Pregnancies and Risk of Preterm Delivery at Split University Hospital Centre: Three-Year Experience. *Acta clinica Croatica* **56**, 640–644 (2017).**

66. Machado, M. *et al.* [Perinatal Outcome in Relation to Chorionicity in Twin Pregnancy]. *Acta medica portuguesa* **30**, 12–16 (2017).**

67. Gilman, S. E., Gardener, H. & Buka, S. L. Maternal smoking during pregnancy and children's cognitive and physical development: a causal risk factor? *American journal of epidemiology* **168**, 522–531 (2008). *

68. Källén, K. The impact of maternal smoking during pregnancy on delivery outcome. *European journal of public health* **11**, 329–333 (2001). *

69. Kwon, J. Y. *et al.* Umbilical arterial blood gas and perinatal outcome in the second twin according to the planned mode of delivery. *International journal of medical sciences* **8**, 643–648 (2011).

70. Usta, I. M. *et al.* Comparison of the perinatal morbidity and mortality of the presenting twin and its co-twin. *Journal of perinatology: official journal of the California Perinatal Association* **22**, 391–396 (2002).**

71. Stein, W., Misselwitz, B. & Schmidt, S. Twin-to-twin delivery time interval: influencing factors and effect on short-term outcome of the second twin. *Acta obstetricia et gynecologica Scandinavica* **87**, 346–353 (2008).**

72. Algeri, P. *et al.* Neonatal hypoxia of the second twin after vaginal delivery of the first twin: what matters? *The journal of maternal-fetal & neonatal medicine: the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians* **32**, 2889–2896 (2019). **

73. Rayburn, W. F., Lavin, J. P. J., Miodovnik, M. & Varner, M. W. Multiple gestation: time interval between delivery of the first and second twins. *Obstetrics and gynecology* **63**, 502–506 (1984). **

74. Schneuber, S. *et al. Twin-to-twin delivery time: neonatal outcome of the second twin. Twin research and human genetics : the official journal of the International Society for Twin Studies* **14**, 573–579 (2011). **

75. Krebs, L. & Langhoff-Roos, J. Breech delivery at term in Denmark, 1982–92: a population-based case-control study. *Paediatric and perinatal epidemiology* **13**, 431–441 (1999). *

76. Krebs, L., Langhoff-Roos, J. & Thorngren-Jerneck, K. Long-term outcome in term breech infants with low Apgar score--a population-based follow-up. *European journal of obstetrics, gynecology, and reproductive biology* **100**, 5–8 (2001). *

77. Vogel, J. P. *et al.* Outcomes of non-vertex second twins, following vertex vaginal delivery of first twin: a secondary analysis of the WHO Global Survey on maternal and perinatal health. *BMC pregnancy and childbirth* **14**, 55 (2014). **

78. Bjelic-Radisic, V. *et al.* Neonatal outcome of second twins depending on presentation and mode of delivery. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **10**, 521–527 (2007). **

79. Schmitz, T. *et al.* Neonatal outcomes of twin pregnancy according to the planned mode of delivery. *Obstetrics and gynecology* **111**, 695–703 (2008). **

80. Jhaveri, R. R. & Nadkarni, T. K. Perinatal Outcome of Second Twin with Respect to Mode of Delivery: An Observational Study. *Journal of clinical and diagnostic research: JCDR* **10**, QC26–QC28 (2016). **

81. Usta, I. M., Rechdan, J. B., Khalil, A. M. & Nassar, A. H. Mode of delivery for vertex-nonvertex twin gestations. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* **88**, 9–14 (2005). **

82. Wenckus, D. J., Gao, W., Kominiarek, M. A. & Wilkins, I. The effects of labor and delivery on maternal and neonatal outcomes in term twins: a retrospective cohort study. *BJOG: an international journal of obstetrics and gynaecology* **121**, 1137–1144 (2014).**

83. Barrett, J. F. R. *et al.* A randomized trial of planned cesarean or vaginal delivery for twin pregnancy. *The New England journal of medicine* **369**, 1295–1305 (2013). **

84. Fox, N. S. *et al.* Active second-stage management in twin pregnancies undergoing planned vaginal delivery in a U.S. population. *Obstetrics and gynecology* **115**, 229–233 (2010).**

85. Sentilhes, L. *et al.* Attempted vaginal versus planned cesarean delivery in 195 breech first twin pregnancies. *Acta obstetricia et gynecologica Scandinavica* **86**, 55–60 (2007).**

86. Ladehoff, P., Pedersen, G. T. & Sørensen, T. Apgar scores in low birth weight infants delivered vaginally and by cesarean section. *Acta obstetricia et gynecologica Scandinavica* **65**, 3–5 (1986). *

87. Stevenson, D. K. *et al.* Sex differences in outcomes of very low birthweight infants: the newborn male disadvantage. *Archives of disease in childhood. Fetal and neonatal edition* **83**, F182–5 (2000). *

88. Zhao, D., Zou, L., Lei, X. & Zhang, Y. Gender Differences in Infant Mortality and Neonatal Morbidity in Mixed-Gender Twins. *Scientific reports* **7**, 8736 (2017).**

89. Bodnar, L. M., Pugh, S. J., Abrams, B., Himes, K. P. & Hutcheon, J. A. Gestational weight gain in twin pregnancies and maternal and child health: a systematic review. *Journal of perinatology: official journal of the California Perinatal Association* **34**, 252–263 (2014).

90. Dubé, J., Dodds, L. & Armson, B. A. Does chorionicity or zygosity predict adverse perinatal outcomes in twins? *American journal of obstetrics and gynecology* **186**, 579–583 (2002).

91. Hogle, K. L., Hutton, E. K., McBrien, K. A., Barrett, J. F. r & Hannah, M. E. Cesarean delivery for twins: a systematic review and meta-analysis. *American journal of obstetrics and gynecology* **188**, 220–227 (2003).

92. Smith, G. C. S., Pell, J. P. & Dobbie, R. Birth order, gestational age, and risk of delivery related perinatal death in twins: retrospective cohort study. *BMJ (Clinical research ed.)* **325**, 1004 (2002).

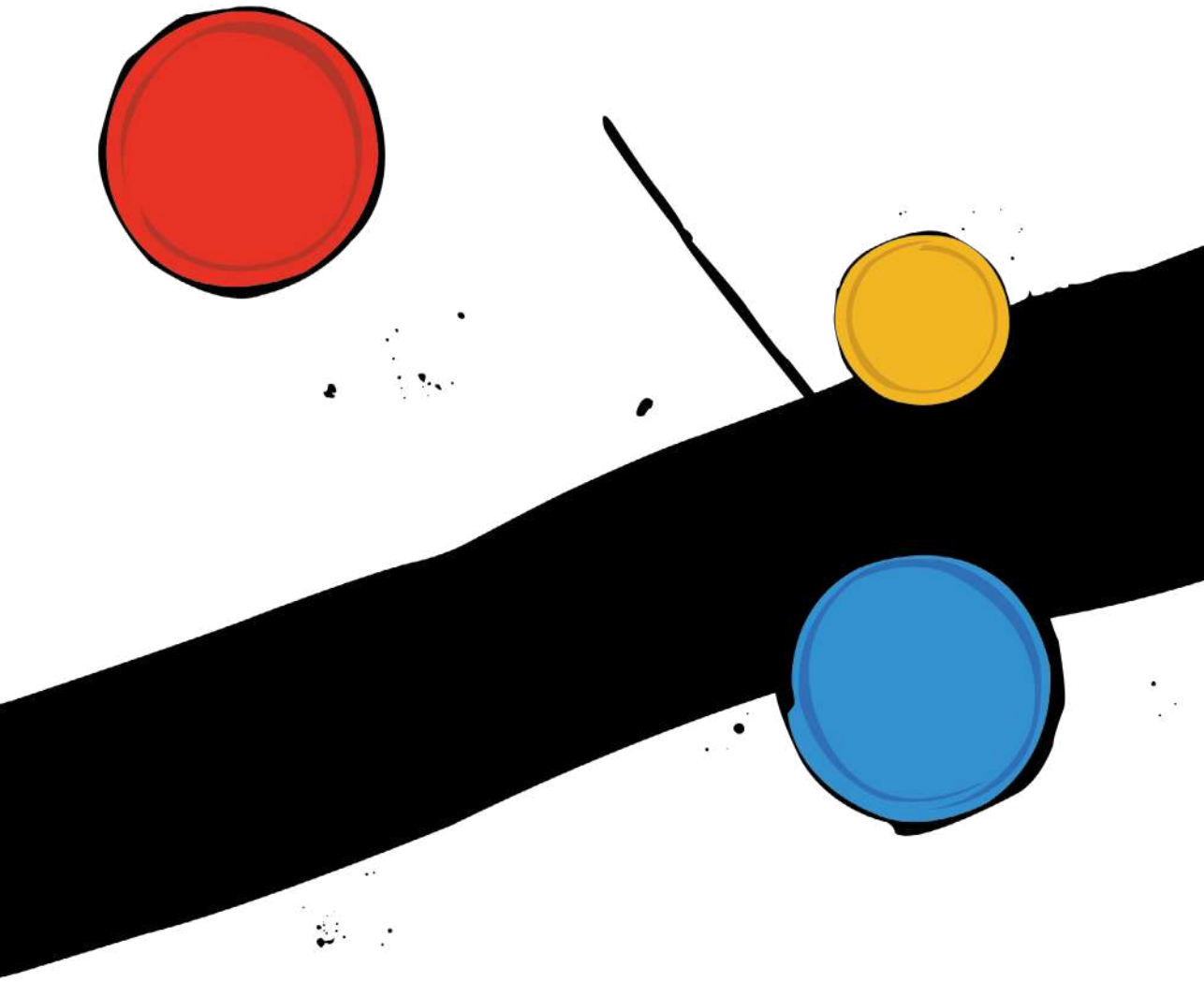
Supplements

The supplementary materials are available online at

https://doi.org/10.1017/thg.2019.24

Chapter 3

Maternal smoking and DNA methylation abnormalities in children at early developmental stages



Abstract

This chapter is based on the review published in the Russian journal *Akusherstvo i Gynekologiya* published from 1930-es and retained as one of the oldest journals in Russian Federation. It discusses the results of studies on the role of DNA methylation during human embryonic development, and the effects of maternal smoking on the epigenetic status of a developing child. The molecular mechanisms mediating the association between maternal smoking during pregnancy and its long-term effects on the development and health of the offspring are the object of active research in medicine and biology. Human genomics studies in recent years have shown that one mechanism may be stable tobacco smoke-induced alterations in DNA methylation that cause concomitant smoking-related developmental and health problems. Maternal smoking during pregnancy has a double effect: firstly, it adversely affects women's health, and secondly, it can lead to irreparable fetal developmental disorders, and affect the health and development of a newborn and the quality of his/her subsequent life.

Published as: Odintsova, V.V., Saifitdinova, A.F., Naumova, O.Y. (2018). Maternal smoking and DNA methylation abnormalities in children at early developmental stages. *Akusherstvo i Ginekologiya* (Russian Federation), (9): 5–12.

Introduction

According to the WHO, tobacco smoking is one of the leading risk factors for morbidity and premature mortality among middle-aged adults. Moreover, for women, it is the second most important risk factor after high blood pressure. While in adults smoking leads to the risk of developing various systemic diseases[1], the effects of smoking during pregnancy can lead to irreparable fetal malformations[2], and can affect the health and development of the newborn, and ultimately quality of his/her life in general. It has been established that smoking during pregnancy acts as an exogenous factor which interferes with the fetus development in a dose-dependent manner[3]. Thus, smoking during pregnancy is known to be associated with respiratory failure, and more pronounced asthma symptoms in childhood, low birth weight, orofacial fissures, sudden sensorineural hearing loss, or sudden deafness in newborns, otitis media, neurobehavioral disorders, etc.[4–10]. A large population-based study of 15,228 pairs of Dutch twins from the Netherlands Twin Register showed a greater effect of maternal smoking during pregnancy compared to paternal smoking for externalizing behaviour, aggression, overactive and withdrawn behaviour. Smoking cessation before pregnancy was associated with less externalizing, overactive behaviour, aggression and oppositional behaviour, but had no effect on internalizing behaviour, anxious depression, or withdrawn behaviour[11].

Traditional research studying the relative influence of genetic and environmental factors on individual variation in health and morbidity indicators tends to focus on the association between the disease, external factors and genotype. Studies conducted over recent decades in the field of human genomics have shown that epigenetic regulation of gene expression, particularly in the early stages of development, plays a significant role in the development of diseases and systemic disorders, along with genomic variation[12]. Epigenetic regulation is attributable to specific mechanisms which control the functional activity of gene promoters (nucleic acid sequences, recognizable by RNA polymerase and transcription factors as a starting point for transcription), and other regulatory regions of the genome and include DNA methylation, histone modifications, chromatin remodelling and RNA interference[13]. *Epigenetics* (from the Greek *epi* – upon, above) focuses on these mechanisms while studying the patterns of heritable changes in gene expression that do not involve changes to the underlying DNA sequence.

This review focuses on one of the main epigenetic mechanisms, DNA methylation, and its role in human embryonic development, with a special emphasis on the impact of tobacco smoking during pregnancy on the child's epigenetic status.

DNA methylation as an epigenetic mechanism

DNA methylation is a mechanism of gene expression regulation and is involved in a number of key processes, including genomic imprinting, X-chromosome inactivation[14], repression of transposable elements[15] and tissue-specific gene expression[16] that underlies the structural and functional diversity of cells and tissues in the body. DNA methylation is a process by which a methyl group (CH_3) is added to cytosine, one of the DNA's nitrogenous bases. This modification predominantly occurs in CpG dinucleotide sites (C – cytosine – which is followed by G – guanine – linked by a phosphate group, p). This site has a pair in the complementary strand of the DNA double helix. This makes it possible

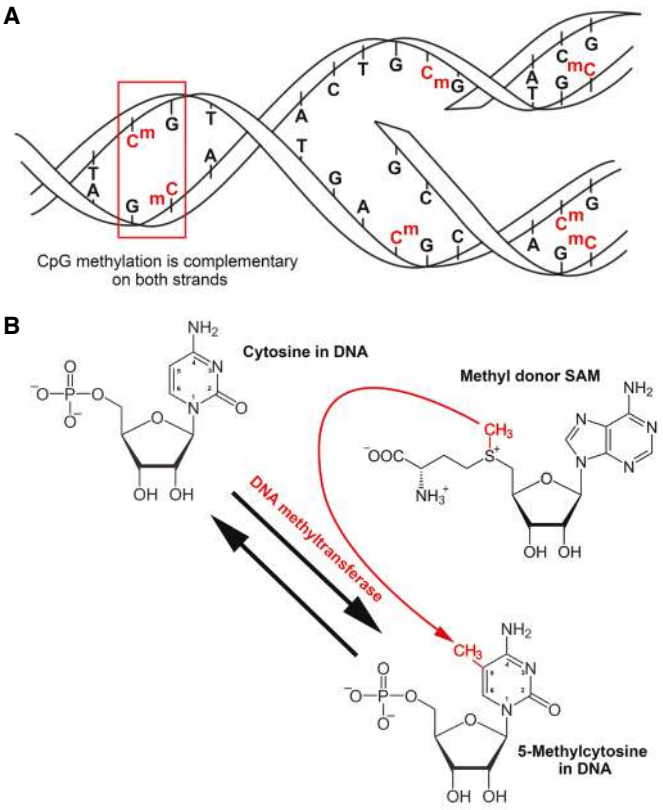


Figure 1. Molecular basis of DNA methylation

to restore methylation on the newly synthesized daughter strand after DNA replication (**Figure 1A**), which, in turn, allows the DNA methylation pattern to be propagated over many cell generations. CpG-rich genomic regions are called CG islands (CGI). About 70% of human genes have a high CpG content in their promoter regions.

DNA methylation is catalysed by a group of enzymes, of which DNA methyltransferases 1 and 3 (DNMT1 and DNMT3) play the major roles. DNMT1 has an intrinsic affinity for hemi-methylated DNA. DNMT1 is responsible for the maintenance of methylation during DNA replication, which is particularly important for those tissues in which active cell division takes place throughout life. The DNMT3 family members interact with unmethylated DNA; they play an important role in establishing a tissue-specific *de novo* methylation pattern at the earliest stages of development[17]. The ratio of methylated and unmethylated cytosines required for normal development is maintained through a balanced methyltransferase activity and demethylation; the process of removal of methyl groups from DNA molecules (**Figure 1B**).

DNA methylation is a reversible process[18]. Whereas the DNA sequence is fixed, except for somatic mutations, methylation is and DNA methylation marks can be removed or added due to environmental exposures, or by epigenetic therapy. DNA methylation is

a multiple stage process controlled by several monitoring systems. DNA demethylation can occur through a number of processes including DNA demethylation mediated by enzymes. In the developing human embryos, the demethylation with an intermediate oxidation state of methylcytosine to hydroxymethylcytosine plays an important role[19]. The proteins of the TET family (i.e., Ten-Eleven Translocation enzymes, whose hyperactivity was first found in the cells with chromosomal translocation t(10;11)(q22;q23) in patients with acute myeloid leukaemia) contribute to the rapid loss of methylation in CpG islands facilitating access to specific gene promoters for transcription factors.

It is known that partly CpG sites are methylated in a tissue-specific manner[16]. The methylation pattern is inherited from one cell generation to next, but cells can adjust it, and activate certain genes in response to external factors. Abnormalities in and failures of this stable and controlled system may lead to development of diseases. For example, an increase in the activity of demethylases has been identified as a contributing factor to the development of cancer. On the contrary, the suppression of their activity results in an increase in methylation level of CpG islands in gene promoters leading to a disruption of their functioning[20].

DNA methylation patterns change with age in response to the influence of acquired experience and environment, lifestyle and socio-emotional factors[21,22], i.e. have a certain plasticity. In a healthy human, methylation and demethylation are under strict control, and facilitate the prompt reaction of the organism to the changing environments, including the internal environment. Having a dynamic nature, DNA methylation is a widely used biomarker in the research aimed to investigate the role of the environment in the development of cancer and other diseases in the context of their molecular aetiology[23–27].

DNA methylation dynamics during embryogenesis

DNA methylation changes in adults can lead to the alteration of the functioning of individual cells and tissues. However, in a developing organism, even a short-term disruption in coordinated epigenetic regulation may impact the organism's development, and may cause irreparable damage. The severity of such damage depends on the magnitude of epigenetic disruptions and the stage of development during which they occurred. Specifically, during gonadal and embryonic development the most rapid and dramatic dynamic changes occur when cell-specific methylation patterns are established via several waves of the global demethylation and de novo methylation [28,29].

Before conception, DNA methylation plays a role in the formation of parental germ cells or gametes. The DNA of primordial germ cells is considerably methylated; when cells migrate to undifferentiated gonads, the methylation level drops. During germ cell maturation, epigenetic information is largely erased as the DNA goes through active demethylation. These processes differ in oogenesis and spermatogenesis, and, as a result, the genetic information from parents is marked differently[30,31].

Immediately after fertilization, the second round of epigenetic reprogramming begins, which is accompanied by global changes in the DNA methylation and histone modifications. In a zygote, the level of DNA methylation decreases first in the paternal pronucleus and then in the maternal pronucleus[19,30,31]. Methylation levels continue to drop

until a morula formation. This process involves both active and passive demethylation mechanisms. At the blastocyst stage, the process of remethylation starts, which provides a blastomere differentiation into the inner cell mass (ICM) and the trophectoderm (TE). After implantation, as a result of *de novo* methylation, epigenetic differences between ICM cells and TE are established. These newly established DNA methylation patterns are preserved during subsequent cell divisions; further differentiation of ICM cells into three germ layers is accompanied by the establishment of the layer-specific DNA methylation patterns through local methylation changes. Both maternal and paternal alleles have equal opportunities to be expressed in any cell of the offspring; for most genes, maternal and paternal alleles are expressed equally. However, several hundred out of approximately 25 thousand human genes are subject to genomic imprinting[32] – a process that causes genes to be expressed in a parent-of-origin-specific manner, and which involves differential methylation of maternal and paternal alleles.

The data on DNA methylation dynamics have been predominantly provided by research in using model organisms[33,34]. Altogether they indicate that there is an early development stage when epigenetic marks are removed, followed by the establishment of *de novo* methylation profiles and chromatin remodelling. In humans, reprogramming of the parental genomes occurs during the first cell divisions. At the morula stage the establishment of new epigenetic signatures begins that is specific for differentiating embryo tissues. Any disruptions of epigenetic reprogramming, including those resulting from the negative impact of external factors, may disturb the ontogenesis program causing aberrant gene expression, which, in turn, may lead to severe pathologies or atrophies, embryonic death or fetal malformations[19,29].

Maternal smoking during pregnancy and offspring DNA methylation

There are numerous publications and literature reviews on the effects of maternal smoking during pregnancy[35–39]. In addition to an indirect impact through the deterioration of the mother's health, direct consequences of maternal smoking on fetal development are widely known in paediatrics[4]. These include the effects of nicotine accumulation in the fetal blood and amniotic fluid (it should be noted that the nicotine concentration in the fetus is usually 15% higher than maternal levels). Nicotine exposure was associated with adverse outcomes in every trimester of pregnancy: from spontaneous abortions during the first trimester to increased risk of premature birth and low birth weight during the last trimester. Nicotine affects birth weight—gestational age and fetal growth rate. Both animal models and human research have revealed that nicotine increases the mother's blood pressure and heart rate, while reducing uteroplacental blood flow. In addition, carbon monoxide in tobacco smoke forms carboxyhaemoglobin, which inhibits the release of oxygen into fetal tissues. Together with the effect of cancerogenic xenobiotics contained in cigarette smoke, this leads to toxication and hypoxia, affecting many systems of the developing fetus, especially the respiratory and nervous systems[4–6]. Consequently, this disrupts normal development of the child, and may lead to congenital brain defects and malformations in other organs, and behavioural impairments manifested in crying for no reason, sleep disturbances, and later, in aggression, and other behavior problems.

Active smoking may disrupt DNA methylation, and the prenatal period, is the most critical developmental stage for establishing an individual's epigenetic status. The first few

comparative studies showed that the genomes of blood cells in the preschool and primary school children with a history of maternal smoking during pregnancy were characterized by changes in the global and gene-specific DNA methylation[40]. Moreover, a study of the association between these changes in DNA methylation and the duration of exposure to tobacco smoke found that the methylation signatures point to stable, rather than short-term effect of maternal smoking during pregnancy[41]. It has been shown that many epigenetic changes associated with maternal smoking have a long-term effect; they are persistent throughout adult life and are detectable around age 40 regardless of active smoking[42].

The studies of the effects of maternal smoking on offspring development utilize two different approaches. The first is a 'retrospective' approach, which examines the long-term epigenetic consequences comparing the cohorts of adults[42] and (young) children[40,43] with a history of in utero exposure to maternal smoking with an age-matched comparison cohort of individuals without a history of maternal smoking during pregnancy. The second approach is based on studying the characteristics of DNA methylation in placental and/or cord blood cells[7,9,41,44–50], and neonatal peripheral blood[51], or fetal tissues. Studies based on fetal tissues are commonly used in the research based on animal models[52]; when applied to humans, such studies are based on the tissues of aborted embryos[45,53–55].

It should be noted that, like research focused on the epigenetic effects of active smoking in adults, the tissue-specific epigenetic response of the fetus to maternal smoking has been barely studied. At the same time, available evidence indicates that such a differentiated response is a reality. Thus, methylation differences for certain genes (*AHRR* and *CYP1A1*) have been found between placental and fetal tissues[48], and the overlap of the sets of genes associated with maternal smoking in blood cells and brain tissues was small53. Another issue is that research implementing fetal tissues is conducted on the epigenetic system, which is extremely dynamic during the prenatal period. The timing of fetal development is crucial because it is not known when exactly during development the specific epigenetic alterations, which are caused by maternal smoking occur and stabilize, and become detectable. Thus, a study on brain tissues from embryos in the second trimester showed that the stage (early or late) of the trimester characterized by rapid brain development rather than maternal smoking was the main factor differentiating individual epigenomes in the studied cohort[53]. In addition, the authors[53] established the effects of maternal smoking on offspring DNA methylation. The greatest effect of prenatal exposure to nicotine was attributed to the epigenetic changes associated with slower neuronal maturation and/or a decreasing number of mature neurons.

One of the first large-scale whole-epigenome studies published by Joubert with co-authors[41,46] was conducted on a cohort of 1,062 newborns (Norwegian Mother and Child Cohort Study). Having analyzed DNA methylation in cord blood cells and assessed the impact of prenatal tobacco exposure based on circulating maternal cotinine (a nicotine metabolite and a stable biomarker of cigarette-smoking), the authors identified epigenetic changes in 10 genes significantly associated with maternal cotinine levels. The most important among them were the genes involved in the xenobiotic-detoxication, namely *AHRR* and *CYP1A1*. It is noteworthy that these two genes have been consistently detected in the association with active smoking in adults. These results were confirmed, and were supported by additional findings in a number of studies[7,43,46,47,50,51].

The Pregnancy and Childhood Epigenetics consortium (PACE) combined data across studies that utilized the same platform for whole-genome DNA methylation profiling—the Infinium HumanMethylation450 array. This consortium conducted a large-scale epigenome-wide meta-analysis using the data on whole-genome DNA methylation in cord blood from 6,685 neonates in 13 birth cohort studies from the US and Europe with a detailed history of maternal smoking during pregnancy[49]. To date, this is the most extensive and comprehensive study of the effect of maternal smoking during pregnancy on DNA methylation in offspring. Using the results from older children cohorts, the researchers investigated the long-term epigenetic alterations triggered by prenatal exposure to nicotine with a correction for the effects of postnatal second-hand tobacco smoke exposure. They also took into account whether the pre-existing smoking history and smoking frequency during pregnancy involved sustained or occasional smoking. The cohort-specific associations between maternal smoking during pregnancy and DNA methylation in offspring were a subject of the meta-analysis. To establish the functional significance of newly identified epigenetic alterations, the authors evaluated the associations between the methylation status and expression level of a number of genes, and performed functional analysis of gene-networks involved in the response to prenatal exposure to nicotine.

The following observations and conclusions followed from the results of this large-scale study[49].

(1) Maternal smoking during pregnancy causes significant changes in the fetal epigenome. This study revealed over 6 thousand CpG sites across the genome, which showed significant changes in methylation status in the children whose mothers smoked during pregnancy, with the *AHRR* gene taking a leading position with respect to the association significance ($1.64x10^{-193}$).

(2) The epigenetic alterations established in newborns are long-lasting epigenetic changes. Thus, although an attenuation of effects was observed in older children, there was a high concordance in the methylation statuses of the newborns and older children with a history of maternal smoking. This concordance was confirmed statistically for 73% (4 out of 6 thousand) of the CpGs.

(3) As expected, any exposure to maternal smoking affects the fetus. Sustained maternal smoking, however, is characterized by a greater effect on epigenetic alterations, that, in turn, may be associated with a higher impact of maternal smoking on the child's health and development.

(4) The fact that such negative health and developmental outcomes exist is well-known, but that they may be partly due to epigenetic alterations driven by maternal smoking is suggested by the results of the functional analysis of the genes, whose methylation changes as a result of prenatal exposure to tobacco smoke. Thus, these genes are predominantly involved in the control of key developmental processes, such as growth and anatomical development, including those specifically related to embryonic morphogenesis (or during embryogenesis), of the nervous system development, cell growth and proliferation[49].

The epigenetic changes as one of the molecular mechanisms linking the effects of prenatal exposure to nicotine with various negative consequences for the child's health and development have been confirmed by a few additional studies. These studies identified the associations between maternal smoking – epigenome – phenotypes. The associated phenotypes include obesity[7], inadequate immune response[9], and the central nervous system developmental disorders[8].

Epigenetics of smoking. Open questions and future research

Smoking-related diseases continue to be a major public health concern, and understanding the mechanisms of the health effects of smoking is an important part of research. Research of the past 10 years has shown that chronic exposure to tobacco smoke is an adverse environmental stimulus that is capable of modifying DNA methylation patterns. These modifications, in turn, may extend the effects of smoking onto gene expression and, eventually, lead to impairments and diseases associated with smoking. It has been found that the alterations in DNA-methylation driven by tobacco smoke can occur at all stages of development: in adult life, as a result of active smoking, and during prenatal development, as a result of maternal smoking. In adulthood, smoking affects methylation patterns, which are already established, and are maintained throughout life and during cell division. On the other hand, prenatal exposure to tobacco smoke happens at a time when DNA methylation is highly dynamic, and methylation patterns evolve. As a result, prenatal effects of tobacco smoke can be larger, having a significant impact on the program of embryonic development.

Tobacco smoke exposure affects specific genomic regions. Thus, identical alterations in methylation patterns of genes involved in chemical detoxication of tobacco smoke components (for example, *AHRR* and *CYP1A1*) are found in smokers' somatic tissues, smoking mothers' placenta, and in fetal tissues. In addition, smoking destabilizes methylation patterns at the global genomic level. It is also likely that the pattern of these alterations in different cells and tissues may vary affecting different molecular cascades. This was observed both in somatic tissues of adult smokers and in differentiating embryonic tissues.

Despite significant progress in research of the epigenetics of smoking, many issues require further research; of those, the following two merit consideration. First, while an association between smoking and methylation disturbances in certain genes has been established, the epigenome-wide response to smoking, in terms of the genome dysregulation at the level of gene systems and gene cascades, remains poorly studied. Second, it has been shown that smoking can change DNA methylation in various tissues, and some of these changes may vary between the tissues[53,56,57]. However, the majority of studies were carried out on peripheral blood cells, which are the most accessible biological material. Thus, the issue of differential epigenomic response to smoking in various tissues and organs remains understudied.

Further research of the effect of smoking on the epigenome, including its impact on DNA methylation, has a high potential to advance our understanding in mediation effects of DNA methylation and causal effects. Firstly, new findings in this field may contribute to understanding the molecular mechanisms of the impact of tobacco smoke on the

body and the development of impairments and health issues associated with smoking. Secondly, the epigenetic alterations resulting in radical changes in the phenotype are particularly interesting because by nature they are either fully or partially reversible. This reversibility potentially allows running a scenario where the 'epigenetic code' can dictate the expression of a particular set of genes, essentially acting as an on/off switch. 'Epigenetic drugs' have already been developed, and used in the treatment of neurological diseases and cancers[58,59]. Regarding the smoking, any epigenetic drug, which could counteract its harmful effects, remains, at present, science fiction. Nevertheless, even today's studies of the influence of certain substances on the epigenetic outcomes of maternal smoking in offspring have borne fruitful results. For instance, a randomized clinical study showed that adding vitamin C to the diet of a smoking pregnant woman helps to normalize DNA methylation in the baby, and to reduce respiratory problems in babies, which is an established effect of maternal smoking[45].

References

1. Fagerström, K. The epidemiology of smoking: health consequences and benefits of cessation. *Drugs* **62 Suppl 2**, 1–9 (2002).
2. Hackshaw, A., Rodeck, C. & Boniface, S. Maternal smoking in pregnancy and birth defects: a systematic review based on 173 687 malformed cases and 11.7 million controls. *Hum. Reprod. Update* **17**, 589–604 (2011).
3. Meberg, A., Sande, H., Foss, O. P. & Stenwig, J. T. Smoking during pregnancy-effects on the fetus and on thiocyanate levels in mother and baby. *Acta Paediatr. Scand.* **68**, 547–552 (1979).
4. Lambers, D. S. & Clark, K. E. The maternal and fetal physiologic effects of nicotine. *Semin. Perinatol.* **20**, 115–126 (1996).
5. Neuman, Å. *et al.* Maternal smoking in pregnancy and asthma in preschool children: a pooled analysis of eight birth cohorts. *Am. J. Respir. Crit. Care Med.* **186**, 1037–1043 (2012).
6. Roza, S. J. *et al.* Effects of maternal smoking in pregnancy on prenatal brain development. The Generation R Study. *Eur. J. Neurosci.* **25**, 611–617 (2007).
7. Janssen, B. G. *et al.* Placental mitochondrial DNA and CYP1A1 gene methylation as molecular signatures for tobacco smoke exposure in pregnant women and the relevance for birth weight. *J. Transl. Med.* **15**, 5 (2017).
8. Banik, A. *et al.* Maternal Factors that Induce Epigenetic Changes Contribute to Neurological Disorders in Offspring. *Genes (Basel).* **8**, (2017).
9. Van Otterdijk, S. D., Binder, A. M. & Michels, K. B. Locus-specific DNA methylation in the placenta is associated with levels of pro-inflammatory proteins in cord blood and they are both independently affected by maternal smoking during pregnancy. *Epigenetics* **12**, 875–885 (2017).
10. Orlebeke, J. F., Boomsma, D. I., Van Baal, G. C. M. & Bleker, O. P. Effect of maternal smoking on birth weight of twins: a study from the Dutch Twin Register. *Early Hum. Dev.* **37**, 161–166 (1994).
11. Dolan, C. *et al.* Testing Causal Effects of Maternal Smoking During Pregnancy on Offspring's Externalizing and Internalizing Behavior. *Behav. Genet.* **46**, (2015).
12. Dolinoy, D. C., Weidman, J. R. & Jirtle, R. L. Epigenetic gene regulation: linking early developmental environment to adult disease. *Reprod. Toxicol.* **23**, 297–307 (2007).
13. Razin, A. CpG methylation, chromatin structure and gene silencing-a three-way connection. *EMBO J.* **17**, 4905–4908 (1998).
14. Escamilla-Del-Arenal, M., da Rocha, S. T. & Heard, E. Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Hum. Genet.* **130**, 307–327 (2011).
15. Fazzari, M. J. & Greally, J. M. Epigenomics: beyond CpG islands. *Nat. Rev. Genet.* **5**, 446–455 (2004).
16. Eckhardt, F. *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* **38**, 1378–1385 (2006).
17. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
18. Ramchandani, S., Bhattacharya, S. K., Cervoni, N. & Szyf, M. DNA methylation is a reversible biological signal. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 6107–6112 (1999).
19. Efimova, O. A. *et al.* Chromosome hydroxymethylation patterns in human zygotes and cleavage-stage embryos. *Reproduction* **149**, 223–233 (2015).
20. Wu, H. *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**, 389–393 (2011).
21. Meaney, M. J. & Szyf, M. Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. *Dialogues Clin. Neurosci.* **7**, 103–123 (2005).

22. Naumova, O. Y. *et al.* Epigenetic Patterns Modulate the Connection Between Developmental Dynamics of Parenting and Offspring Psychosocial Adjustment. *Child Dev.* **87**, 98–110 (2016).
23. Calvanese, V., Lara, E., Kahn, A. & Fraga, M. F. The role of epigenetics in aging and age-related diseases. *Ageing Res. Rev.* **8**, 268–276 (2009).
24. Landgrave-Gómez, J., Mercado-Gómez, O. & Guevara-Guzmán, R. Epigenetic mechanisms in neurological and neurodegenerative diseases. *Front. Cell. Neurosci.* **9**, 58 (2015).
25. Robertson, K. D. DNA methylation and human disease. *Nat. Rev. Genet.* **6**, 597–610 (2005).
26. Jirtle, R. L. & Skinner, M. K. Environmental epigenomics and disease susceptibility. *Nat. Rev. Genet.* **8**, 253–262 (2007).
27. Rakyan, V. K., Down, T. A., Balding, D. J. & Beck, S. Epigenome-wide association studies for common human diseases. *Nat. Rev. Genet.* **12**, 529–541 (2011).
28. Chavez, L. *et al.* Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. *Genome Res.* **20**, 1441–1450 (2010).
29. Haaf, T. Methylation dynamics in the early mammalian embryo: implications of genome reprogramming defects for development. *Curr. Top. Microbiol. Immunol.* **310**, 13–22 (2006).
30. Dean, W. DNA methylation and demethylation: a pathway to gametogenesis and development. Mol. *Reprod. Dev.* **81**, 113–125 (2014).
31. Reik, W. & Surani, M. A. Germline and Pluripotent Stem Cells. *Cold Spring Harb. Perspect. Biol.* **7**, (2015).
32. Barlow, D. P. & Bartolomei, M. S. Genomic imprinting in mammals. *Cold Spring Harb. Perspect. Biol.* **6**, (2014).
33. Smith, Z. D. *et al.* DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611–615 (2014).
34. Messerschmidt, D. M., Knowles, B. B. & Solter, D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* **28**, 812–828 (2014).
35. Knopik, V. S., Marceau, K., Bidwell, L. C. & Rolan, E. Prenatal substance exposure and offspring development: Does DNA methylation play a role? *Neurotoxicol Teratol* **71**, 50–63 (2019).
36. Banderali, G. *et al.* Short and long term health effects of parental tobacco smoking during pregnancy and lactation: a descriptive review. J. *Transl. Med.* **13**, 327 (2015).
37. Anderson, T. M. *et al.* Maternal Smoking Before and During Pregnancy and the Risk of Sudden Unexpected Infant Death. *Pediatrics* **143**, (2019).
38. Berlin, I. & Oncken, C. Maternal Smoking During Pregnancy and Negative Health Outcomes in the Offspring. *Nicotine Tob. Res.* **20**, 663–664 (2018).
39. Nakamura, A., François, O. & Lepeule, J. Epigenetic Alterations of Maternal Tobacco Smoking during Pregnancy: A Narrative Review. *Int. J. Environ. Res. Public Health* **18**, (2021).
40. Breton, C. V. *et al.* Prenatal Tobacco Smoke Exposure Affects Global and Gene-specific DNA Methylation. *Am. J. Respir. Crit. Care Med.* **180**, 462–467 (2009).
41. Joubert, B. R. *et al.* 450K Epigenome-Wide Scan Identifies Differential DNA Methylation in Newborns Related to Maternal Smoking during Pregnancy. *Environ. Health Perspect.* **120**, 1425–1431 (2012).
42. Tehranifar, P. *et al.* Maternal cigarette smoking during pregnancy and offspring DNA methylation in midlife. *Epigenetics* **13**, 129–134 (2018).
43. Rzehak, P. *et al.* Maternal Smoking during Pregnancy and DNA-Methylation in Children at Age 5.5 Years: Epigenome-Wide-Analysis in the European Childhood Obesity Project (CHOP)-Study. *PLoS One* **11**, e0155554 (2016).
44. Sengupta, S. M., Smith, A. K., Grizenko, N. & Joober, R. Locus-specific DNA methylation changes and phenotypic variability in children with attention-deficit hyperactivity disorder. *Psychiatry Res.* **256**, 298–304 (2017).

45. Shorey-Kendrick, L. E. *et al.* Vitamin C Prevents Offspring DNA Methylation Changes Associated with Maternal Smoking in Pregnancy. *Am. J. Respir. Crit. Care Med.* **196**, 745–755 (2017).

46. Rotroff, D. M. *et al.* Maternal smoking impacts key biological pathways in newborns through epigenetic modification in Utero. *BMC Genomics* **17**, 976 (2016).

47. Morales, E. *et al.* Genome-wide DNA methylation study in human placenta identifies novel loci associated with maternal smoking during pregnancy. *Int. J. Epidemiol.* **45**, 1644–1655 (2016).

48. Suter, M. *et al.* Maternal tobacco use modestly alters correlated epigenome-wide placental DNA methylation and gene expression. *Epigenetics* **6**, 1284–1294 (2011).

49. Joubert, B. R. *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am. J. Hum. Genet.* **98**, 680–696 (2016).

50. Richmond, R. C. *et al.* Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum. Mol. Genet.* **24**, 2201–2217 (2015).

51. Markunas, C. A. *et al.* Identification of DNA Methylation Changes in Newborns Related to Maternal Smoking during Pregnancy. *Environ. Health Perspect.* **122**, 1147–1153 (2014).

52. Meyer, K. F. *et al.* The fetal programming effect of prenatal smoking on Igf1r and Igf1 methylation is organ- and sex-specific. *Epigenetics* **12**, 1076–1091 (2017).

53. Chatterton, Z. *et al.* In utero exposure to maternal smoking is associated with DNA methylation alterations and reduced neuronal content in the developing fetal brain. *Epigenetics Chromatin* **10**, 4 (2017).

54. Fa, S. *et al.* Assessment of global DNA methylation in the first trimester fetal tissues exposed to maternal cigarette smoking. *Clin. Epigenetics* **8**, 128 (2016).

55. Fa, S. *et al.* Changes in first trimester fetal CYP1A1 and AHRR DNA methylation and mRNA expression in response to exposure to maternal cigarette smoking. *Environ. Toxicol. Pharmacol.* **57**, 19–27 (2018).

56. Peters, I. *et al.* Adiposity and age are statistically related to enhanced RASSF1A tumor suppressor gene promoter methylation in normal autopsy kidney tissue. *Cancer Epidemiol. biomarkers Prev. a Publ. Am. Assoc. Cancer Res. cosponsored by Am. Soc. Prev. Oncol.* **16**, 2526–2532 (2007).

57. Satta, R. *et al.* Nicotine decreases DNA methyltransferase 1 expression and glutamic acid decarboxylase 67 promoter methylation in GABAergic interneurons. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16356–16361 (2008).

58. Mann, B. S., Johnson, J. R., Cohen, M. H., Justice, R. & Pazdur, R. FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *Oncologist* **12**, 1247–1252 (2007).

59. Mikaelsson, M. A. & Miller, C. A. The path to epigenetic treatment of memory disorders. *Neurobiol. Learn. Mem.* **96**, 13–18 (2011).

Chapter 4

DNA methylation signatures of breastfeeding in buccal cells collected in mid-childhood

Abstract

Breastfeeding has long-term benefits for children, which may be mediated via the epigenome. This pathway has been hypothesized, but the number of empirical studies in humans is small and mostly done in peripheral blood as the DNA source. We performed an epigenome-wide association study (EWAS) in buccal cells collected around age 9 (mean=9.5) from 1,006 twins recruited by the Netherlands Twin Register (NTR). Age-stratified analysis examined if effects attenuate with age (median split at 10 years; $n_{<10}$ = 517, mean age=7.9; $n_{>10}$ = 489, mean age=11.2). We performed replication analyses in two independent cohorts from NTR (buccal cells) and ALSPAC (peripheral blood), and tested loci previously associated with breastfeeding in epigenetic studies. Genome-wide DNA methylation was assessed with the Illumina Infinium MethylationEPIC BeadChip in NTR and with the HumanMethylation450 Bead Chip in ALSPAC. The duration of breastfeeding was dichotomized ('never' vs 'ever'). In the total sample, no robustly associated epigenome-wide significant CpGs were identified (α=6.34x10^{-08}). In the sub-group of children younger than 10 years, four significant CpGs were associated with breastfeeding, after adjusting for child and maternal characteristics. In children older than 10 years methylation differences at these CpGs were smaller and non-significant. The findings did not replicate in the NTR sample (*n*=98; mean age=7.5 years), and no nearby sites were associated with breastfeeding in the ALSPAC study (*n*=938; mean age=7.4). Of the CpG sites previously reported in the literature, three were associated with breastfeeding in children younger than 10 years, showing that these CpGs are associated with breastfeeding in buccal and blood cells. Our study is the first to show that breastfeeding is associated with epigenetic variation in buccal cells in children. Further studies are needed to investigate if methylation differences at these loci are caused by breastfeeding or by other unmeasured confounders and what mechanism drives changes in associations with age.

Published as: Odintsova, V. V., Hagenbeek, F. A., Suderman, M., Caramaschi, D., van Beijsterveldt, C., Kallsen, N. A., Ehli, E. A., Davies, G. E., Sukhikh, G. T., Fanos, V., Relton, C., Bartels, M., Boomsma, D. I., van Dongen, J. (2019). DNA methylation signatures of breastfeeding in buccal cells collected in mid-childhood. *Nutrients,* 11(11): 2804.

Introduction

Early life environmental influences are associated with development and disease. One mechanism that has been hypothesized to account for the association is through "epigenetic programming" of genes, some of which may act to confer plasticity to developmental processes[1,2]. Nutrition in early life may play a crucial role in modulating gene expression[3,4]. It has been hypothesized that nutrition-induced epigenetic variation may result in different development trajectories and may be associated with metabolic and immune development during critical periods of early life[5–8]. Epigenetic mechanisms are a key element in understanding the developmental origins of later life disease risk[9–11]. One of the best studied epigenetic mechanisms is DNA methylation, the modification of a cytosine base, usually at CpG dinucleotides, with a methyl group, which regulates gene expression and seems to be sensitive to nutrition, already during the prenatal period[12–17].

In humans, prenatal maternal famine has been associated with long-term DNA methylation changes that are still observable in middle-aged individuals[18,19]. The early postnatal period is also believed to be a critical period at which permanent long-term changes may be induced by environmental exposures that affect the child's susceptibility to chronic disease[20], however, whether similar long-term changes in the human epigenome may be induced in this period is unexplored. The first nutrition, including breastfeeding, has long-term effects on children[3]. Human breast milk has a unique composition that differs from other lactating animals and that is quite impossible to reproduce in artificial production; in fact, breast milk contains a unique mixture of microorganisms (the microbiome)[21], metabolites[22], multipotent stem cells[23], growth factors[24] and other components that render it unique and individualized for each newborn[25]. Moreover, breast milk varies its composition according to lactation period, circadian rhythm, and even varies from the start till the end of one feeding[26]. The benefits of breastfeeding on the health of the child are widely described[27–29] and may involve transmission of nutrients, hormones, and antibodies from mother to child[24,30,31]. It also include the process of interaction and attachment of a child to their mother, although the effect of bonding may also be achieved through formula feeding[32,33].

Balanced newborn feeding is the basis for adequate growth and development in childhood and beyond[25,28,34]. Breast milk is important for sensory, neurological and cognitive development[35–37], especially in preterm infants[38], but effects on cognition are confounded by maternal education[39]. The association between breastfeeding duration and lower risk of infectious diseases, obesity, cancer, coronary heart disease, some allergies, autoimmune disease, diabetes mellitus, inflammatory bowel diseases, and metabolic syndrome at later age has been widely studied[27,40–49]. Even though a protective effect of breastfeeding in hypertension, diabetes, obesity, and metabolic syndrome was not evident in some large studies[30,50], the benefits of breastfeeding are well recognized, and the understanding of the biological mechanisms of its influence is of interest. Previous studies on the association between breastfeeding and child outcomes have used a variety of definitions of breastfeeding, including: never vs. ever breastfeeding, breastfeeding duration, which has been assessed as a continuous measure in months or weeks, or as a categorical variable (e.g. long vs. short). Some studies have examined the percentage of breastfed meals, exclusive breastfeeding duration (when a child is exclusively breastfed

until formula feed and/or solid foods are introduced). In meta-analysis by Victora et al.[27] breastfeeding never vs. ever was associated with a reduction in sudden infant deaths, a reduction of acute otitis during the first years of life, a protection against allergic rhinitis in children < 5 years, a higher child IQ of about 3 IQ points. More versus less breastfeeding was associated with major protection against diarrhea morbidity, reduction in severe respiratory infection, and effects on deciduous teeth. Exclusive breastfeeding was associated with strong protective result against infectious disease and allergic rhinitis in children <5 years. A dose-response association with duration of breastfeeding was found for higher IQ and decreased risk for overweight and obesity.

Epigenome-wide association studies (EWASs) can offer new insights to DNA-environment interactions in determining child development and health[51]. To our knowledge seven association studies (2 candidate gene studies, 2 EWASs with breastfeeding as covariate and 3 EWASs of breastfeeding) have been performed on breastfeeding and DNA methylation in humans to date (see **Supplemental Table S1**). A candidate gene study of Obermann-Borst et al. found a negative association between the duration of breastfeeding (in categories) and methylation level in blood cells from 120 children (mean age 1.4 years) at seven CpG sites in the promoter of the *LEP* gene; a hormone that regulates energy homeostasis[52]. A suggestive positive association of the methylation level of 2201 CpGs and a negative association of the methylation level of 2075 CpGs with the duration of breastfeeding (continuous measure in weeks) were reported in blood samples from 37 infants (mean age 25.7 months)[53]. These CpGs were annotated to genes predominantly involved in the control of cell signaling systems, the development of anatomical structures and cells and the development and function of the immune system and the central nervous system[53]. The impact of breastfeeding duration (continuous measure in months) on DNA methylation patterns in 200 children (mean age 11.6 years) was suggested in a study of asthma[54]. An EWAS of Sherwood et al on exclusive breastfeeding supported Obermann-Borst et al. findings, at a later stage in childhood (10 years, *n*=297) but not in young adulthood (18 years, *n*=305)[55]. This suggests that methylation changes induced by breastfeeding may change with time and may be more evident at an early age. Similarly, it has been observed that associations between DNA methylation and maternal smoking and birthweight attenuate during childhood[56,57]. Nevertheless, a long-lasting modulating effect of breastfeeding (continuous measure in weeks) on the effects of methylation quantitative trait loci (mQTLs) for CpG sites at the 17q21 locus, where the *IL4R* (interleukin-4) gene is located, has been suggested at age 18 (*n*=245)[58]. Not having been breastfed has been associated with an increase in methylation of the promoter of the tumor suppressor gene *CDKN2A* (cyclin dependent kinase inhibitor 2A) in premenopausal breast tumors of 639 women (mean age of 57.6 years)[59]. In a more recent EWAS study, breastfeeding (dichotomized as never vs. ever) was associated with changes in the *TTC34* gene at age 7 (*n*=640), which were still evident in adolescence (*n*=709)[60]. These previous epigenetic studies of breastfeeding were often conducted in relatively small samples (average sample size=307, range [37 – 640]). In all studies, DNA was extracted from peripheral blood[52–55,58], or from tumor tissues in adults[59].

We aimed to carry out an EWAS of breastfeeding in 1,006 children around 9 years of age recruited by the Netherlands Twin Register (NTR) based on buccal cell DNA and a replication analysis of loci previously associated with breastfeeding in aforementioned

epigenetic studies. Buccal samples typically consist of a large proportion of epithelial cells, which might serve as a surrogate tissue for other ectodermal tissues, including the brain[61,62]. Buccal samples also consist of a smaller proportion of leukocytes[63]. To date, few EWASs have been performed on buccal DNA. As some studies have suggested that effects of early life exposures, including breastfeeding[55–58], may fade away during childhood, we also performed an EWAS in younger children (age < 10 years; where 10 corresponds to the median age of the sample) and compared effect sizes in this group with effect sizes in children older than 10 years. We applied median split of the sample by age to achieve equal sample sizes in both groups. We hypothesized that if effects of breastfeeding attenuate with age, associations would be strongest in the younger age group. We performed replication in an independent buccal-cells DNA methylation dataset from the NTR (*n*=98) and in a blood-DNA-methylation dataset from the Avon Longitudinal Study of Parents and Children (ALSPAC) (*n*=938). We also examined the correlation between methylation levels of twins for the significant CpGs associated with breastfeeding. We hypothesized that the equal exposure to breastfeeding of co-twins should cause resemblance in their methylation profiles.

Materials and Methods

Overview

We carried out an EWAS in NTR in 1,006 children from 496 complete pairs and 14 twins from incomplete pairs with DNA methylation in buccal cells, testing 787,711 methylation sites (α=0.05/787771= $6.34x10^{-08}$). The EWAS analyses in different age groups were performed following median-split of the sample by age (age < 10 years: *n*= 517, age range [5-9]; age >= 10 years: n = 489, age range [10-12[). Two models were applied with different covariates (model 1="basic model" and model 2="adjusted model") (see **Supplementary Table S2**). Epigenome-wide significant results were taken forward, after checking effects of outliers, for replication analysis in an independent sample, consisting of 98 NTR children with DNA methylation in buccal cells and in 938 ALSPAC children with DNA methylation in peripheral blood cells. Lastly, we performed a follow-up analysis of the results of previous studies (3,859 CpGs) in the discovery cohort (NTR) in the total sample (*n*=1,006) and the younger age group (*n*=517). A flowchart of the analyses is provided in **Figure 1**.

Subjects and Samples

Discovery study

The subjects were enrolled in the Netherlands Twin register (NTR)[64] a few weeks to months after birth. Informed consent was obtained from parents. Data on breastfeeding collected around the age of 2 years of the children and good quality DNA methylation data around 9 years of age (mean=9.5, SD=1.89, range [5-12]) were available for 1,006 children. The dataset included 51.6% girls, and 86% monozygotic twins. This study is embedded in a larger project on childhood aggression (ACTION)[65,66]. From the population based NTR, the ACTION study we identified twins who at least once scored higher or lower on a sum score for aggression[67,68]. The dataset included 51.6% girls, and 86% monozygotic twins.

NTR replication study

An independent group of children from the NTR for whom EPIC array methylation data were previously described63 were included as replication study. This cohort was also embedded in the ACTION project and comprised of 98 monozygotic twins with available data on breastfeeding duration and DNA methylation from buccal swabs (mean age = 7.5 years, SD = 2.4, age range [1-10]). NTR studies were approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Centre, Amsterdam, an Institutional Review Board certified by the U.S. Office of Human Research Protections (IRB number IRB00002991 under Federal-wide Assurance-FWA00017598; IRB/institute codes, NTR 03-180).

ALSPAC replication study

Data came from the Avon Longitudinal Study of Parents and Children (ALSPAC)[69,70], a population-based birth cohort. All pregnant women living in the geographical area of Avon (UK) with expected delivery date between 1 April 1991 and 31 December 1992 were invited to participate. Approximately 85% of the eligible population was enrolled, totaling 14,541 pregnant women who gave informed and written consent. The study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Of these, 938 children had information on breastfeeding duration and DNA methylation from peripheral blood cells measured on the Infinium HumanMethylation450 BeadChip (mean age =7.4, SD =0.13, age range [7.1-8.8]) within the ARIES project[71].

Phenotype data

In the NTR, breastfeeding was assessed in a questionnaire sent to mothers 2 years after the twins were born. Mothers were asked about breastfeeding separately for the first and second born twin. There were six answer categories: 'no', 'less than 2 weeks', '2–6 weeks', '6 weeks to 3 months', '3–6 months', and 'more than 6 months'. For the main analyses breastfeeding as exposure was recoded into 2 categories: 'ever' and 'never'. No information on exclusive and mixed breastfeeding was available. Duration of breastfeeding was used in secondary analysis, based on the 6 categories from the original questionnaire coded from 0 (no) to 5 (more than 6 months). Socio-economic status (SES) was determined in two ways (depending on the version of the survey): 1) SES was obtained from a full description of the occupation of the parents, and was subsequently coded according to the Standard Classification of Occupations[72]; 2) SES was obtained by the EPG-classification scheme[73], combined with information on parental education. Self-reported maternal pre-pregnancy weight just before pregnancy (in kilograms) divided by the square of height (in meters) was used to obtain maternal pre-pregnancy body mass index BMI *(weight/ height^2).* Maternal smoking during pregnancy was reported by mothers for three trimesters of pregnancy and was coded as non-smoking if the mother did not smoke during the entire pregnancy and smoking if the mother smoked at least during one trimester[74]. Mode of conception was classified in three groups: naturally conceived, conceived through stimulation, and

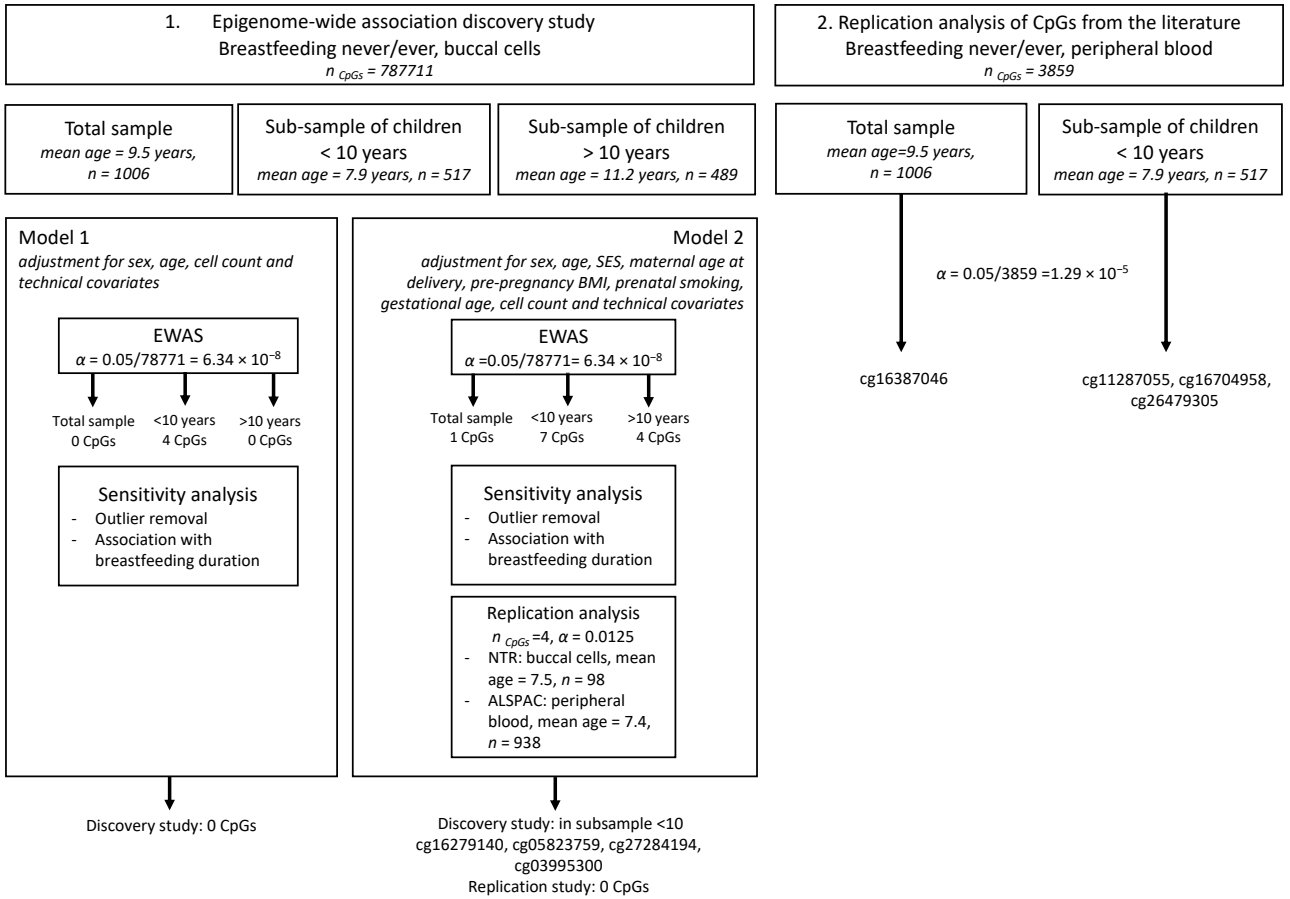


Figure 1. Flowchart of analyses

conceived via in vitro fertilization (IVF) or intracytoplasmic sperm injection (ICSI)[75]. Mode of delivery was assessed as vaginal delivery, caesarean section (planned and urgent), and vaginal delivery with urgent intervention with forceps or vacuum extraction. Apgar scores at 1st and 5th minute were presented in 3 conventional categories 0-6 (low), 7-9 (intermediate), 10 (high)[76]. Gestational age, birthweight, parental age at birth were categorized in groups for descriptive statistics and treated as continuous z-scores in the analyses.

ALSPAC replication study

Breastfeeding was assessed via a questionnaire sent to mothers when the study children were approximately 6 months old. Breastfeeding was coded as 'ever' or 'never'. Information on sample characteristics and covariates was obtained by questionnaires completed by the mother during pregnancy and from medical birth records. Socio-economic status was determined by highest level of maternal education (grouped as follows: certificate of secondary education or not, vocational, O-level, A-level, university degree). Maternal characteristics included age at birth, pre-pregnancy height and weight, and smoking during pregnancy (any or none). Gestational characteristics included caesarean delivery, gestational age and birthweight.

DNA sample collection

NTR (discovery and replication study)

DNA samples were collected from buccal swabs, as described previously[77]. In short, 16 cotton mouth swabs were individually rubbed against the inside of the cheek on 2 days (morning and evening) and placed in four separate 15 mL conical tubes containing 0.5 ml STE buffer (100 mM sodium chloride, 10 mM Tris hydrochloride (pH 8.0) and 10 mM ethylenediaminetetraacetic acid) with proteinase K (0.1 mg/mL) and sodium dodecyl sulfate (SDS) (0.5%) per swab. Individuals were asked to refrain from eating or drinking 1 hour prior to sampling. High molecular weight genomic DNA was extracted by standard DNA extraction techniques. The DNA samples were quantified using the Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific, Waltham, MA, USA).

ALSPAC replication study

ALSPAC blood-based DNA methylation profiles were generated at age 7 as part of the Accessible Resource for Integrated Epigenomics Studies (ARIES)[71], a subsample of approximately 1000 mother-child pairs from the ALSPAC study.

DNA-methylation measurements

NTR discovery study

Genome-wide DNA methylation in buccal cells was assessed with the Infinium MethylationEPIC BeadChip (Illumina, San Diego, CA, USA) 78 by the Human Genotyping facility (HugeF) of ErasmusMC, the Netherlands (http://www.glimdna.org/). 500 ng of genomic DNA from buccal swabs were bisulfite treated using the ZymoResearch EZ DNA Methylation kit (Zymo Research Corp, Irvine, CA, USA). Quality control of the methylation data are described in detail elsewhere. In brief, quality control (QC) and normalization of

the methylation data were performed using a pipeline developed by the Biobank-based Integrative Omics Study (BIOS) consortium[79], which includes sample quality control using the R package MethylAid[80], and probe filtering and functional normalization as implemented in the R package DNAmArray. We previously successfully applied this pipeline in a pilot study of EPIC array data from buccal samples[63], which is used as replication sample in the current paper. MethylAid was applied with the default EPIC array-specific quality filter thresholds for EPIC arrays. The R package omicsPrint[81] was used to call genotypes based on methylation probes and to verify sample relationships based on those Single Nucleotide Polymorphisms (SNPs) (e.g. zygosity of twins and samples from the same individual). We checked for sample mismatches between methylation data and genotype data by computing the correlation between SNP genotypes called by omicsPrint based on methylation probes and genotypes based on genome-wide SNP arrays. DNAmArray and meffil[82] were used to identify sex mismatches (both packages identified the same mismatches).

Functional normalization was performed based on five control probe principal components (PCs). The following probe filters were applied: probes were set to missing (NA) in a sample if they had an intensity value of exactly zero, detection *P*-value > 0.01, or bead count < 3. Probes were excluded from all samples if they mapped to multiple locations in the genome, if they overlapped with a SNP or Insertion/Deletion (INDEL), or if they had a success rate < 0.95 across samples. Annotations of ambiguous mapping probes (based on an overlap of at least 47 bases per probe) and probes where genetic variants (SNPs or INDELS) with a minor allele frequency > 0.01 in Europeans overlap with the targeted CpG or single base extension site (SBE) were obtained from Pidsley et al.[83]. After probe filtering, the success rate of probes for each sample was checked: all samples had a success rate above 0.95 after removal of low-performing samples detected by MethylAid. Only autosomal methylation sites were analyzed, leaving 787,711 out of 865,859 sites for analysis.

NTR replication study

Genome-wide DNA methylation in buccal cells was assessed with the Infinium MethylationEPIC BeadChip (Illumina, San Diego, CA, USA)[84] by the Avera Institute for Human Genetics. Quality control, processing and normalization of the data were performed with the same pipeline as for the NTR discovery cohort and has been described in detail previously[63].

ALSPAC replication study

DNA methylation wet laboratory procedures, preprocessing analyses, and quality control were performed at the University of Bristol as previously described[71]. DNA methylation outliers were identified as those three times the inter-quartile range from the nearest of the first and third quartiles. Outliers were replaced with missing values.

Cellular proportions

NTR discovery and replication study

Cellular proportions were predicted with Hierarchical Epigenetic Dissection of Intra-Sample-Heterogeneity (HepiDISH) with the RPC method (reduced partial correlation),

as described by Zheng et al.[85] and implemented in the R package EpiDISH. HepiDISH is a cell-type deconvolution algorithm that was specifically developed for estimating cellular proportions in epithelial tissues based on genome-wide methylation profiles and makes use of reference DNA methylation data from epithelial cells, fibroblast and seven leukocyte subtypes. This was applied to the data after data QC and normalization.

ALSPAC replication study

Cell count estimates were estimated from DNA methylation profiles using a deconvolution algorithm[86] and included in statistical models to adjust for cell count heterogeneity.

Data analyses

Associations between breastfeeding and pre- and perinatal factors

The association between covariates and breastfeeding was tested in the discovery cohort using a generalized estimating equations (GEE) model accounting for the correlation structure within families. Six breastfeeding duration categories were included as the continuous outcome variable. Predictors consisted of variables previously described as covariates in EWASs of maternal smoking and birth weight[56,57], i.e. parental SES, maternal smoking during pregnancy (yes/no), gestational age (z-scores), maternal age at birth (z-score), maternal pre-pregnancy BMI (z-scores), cell counts of epithelial cells and natural killer (NK) cells, child's sex, and child's age at DNA-methylation (see **Supplemental Table S3**). This analysis was performed in SPSS version 25.

EWAS

Discovery study

The association between DNA methylation level and breastfeeding was tested using a generalized estimating equations (GEE) model accounting for the correlation structure within families with DNA methylation β-value as the outcome variable (see **Supplemental Table S2**). All analyses in NTR were performed with GEE models, which were fitted with the R package 'gee'. The following settings were used: Gaussian link function (for continuous outcome variables), 100 iterations, and the 'exchangeable' option to account for the correlation structure within families. To examine and adjust (where applicable) for inflation of test statistics, the R package bacon was used[87]. We first fitted a basic model with the following predictors: breastfeeding, sex, age at sample collection, percentages of epithelial and natural killer cells, EPIC array row and bisulfite sample plate (using dummy coding). For the primary EWAS, the breastfeeding outcome was dichotomized (yes/no). Second, we fitted a model in which we adjusted for additional covariates (SES, maternal smoking, mother's age at birth, mother's BMI at pregnancy, gestational age). These covariates were selected on the base of adjustments done in recent meta-analyses of maternal smoking and birthweight[56,57]. We also considered effects of duration of breastfeeding by evaluating the same model with the original six categories (not-breastfed; <2 weeks; 2-6 weeks; 6 weeks – 3 months; > 6 months). Thirdly, we repeated the EWAS in children younger than 10 years and older than 10 years. To this end, we

TABLE 1.
Early life characteristics and breastfeeding in the NTR discovery sample (*n*=1,006)

	Breastfeeding Never (*n*=265)		Breastfeeding Ever (*n*=741)		Total	
	N	**%**	***N***	**%**	***N***	**%**
Sex						
male	138	52.1%	349	47.1%	487	48.4%
female	127	47.9%	392	52.9%	519	51.6%
Zygosity						
Monozygotic (MZ)	226	85.3%	613	82.7%	839	83.4%
Dizygotic (DZ)	39	14.7%	128	17.3%	167	16.6%
Chorionicity						
MCMA	4	3.5%	21	7.3%	25	6.2%
MCDA	61	53.0%	144	50.3%	205	51.1%
DCDA	50	43.5%	121	42.3%	171	42.6%
Gestational age (weeks)						
Mean (SD)	36.2	(22.2)	35.8	(25.9)	35.9	(25.1)
<=32	12	4.8%	61	8.4%	73	7.5%
33-36	128	51.4%	359	49.4%	487	49.9%
>=37	109	43.8%	306	42.1%	415	42.6%
Mother's age at birth (years)						
Mean (SD)	31.9	(4.5)	31.2	(4.2)	31.4	(4.3)
19-29	76	29.0%	288	39.0%	364	36.4%
30-39	175	66.8%	435	58.9%	610	61.0%
>40	11	4.2%	15	2.0%	26	2.6%
Mother's BMI before pregnancy						
Mean (SD)	24.3	(4.0)	24.2	(4.11)	24.2	(4.1)
<25	149	61.3%	470	66.1%	619	64.9%
25-29	70	28.8%	169	23.8%	239	25.1%
30-39	24	9.9%	65	9.1%	89	9.3%
>40	0	0.0%	7	1.0%	7	0.7%
Father's age at birth (years)						
Mean (SD)	33.2	(4.4)	33.9	(5.4)	33.7	(5.2)
20-29	53	22.0%	146	20.3%	199	20.8%
30-39	163	67.6%	482	67.1%	645	67.3%
>40	25	10.4%	90	12.5%	115	12.0%
Mode of conception						
naturally	227	92.7%	623	86.5%	850	88.1%
stimulated	4	1.6%	26	3.6%	30	3.1%
IVF/ICSI	14	5.7%	71	9.9%	85	8.8%
Maternal smoking						
no smoking	205	86.1%	631	92.9%	836	91.2%
smoking	33	13.9%	48	7.1%	81	8.8%
Parental SES						
low skill level	0	0.0%	8	1.2%	8	0.9%
lower secondary educational level	30	11.9%	41	6.1%	71	7.6%
upper secondary education level	99	39.3%	203	30.0%	302	32.5%
higher vocational level	89	35.3%	234	34.6%	323	34.8%
scientific level	34	13.5%	191	28.2%	225	24.2%

Table 1 (continued)

	Breastfeeding Never (*n*=265)		Breastfeeding Ever (*n*=741)		Total	
	N	**%**	***N***	**%**	***N***	**%**
Mode of delivery	141	56.6%	416	57.1%	557	57.0%
vaginal	43	17.3%	97	13.3%	140	14.3%
caesarean planned urgent intervention (forceps, vacuum extraction)	20	8.0%	75	10.3%	95	9.7%
urgent caesarean section	45	18.1%	140	19.2%	185	18.9%
Birth weight	2435.7	(444.8)	2394.6	(558)	2405	(531.7)
Mean (SD)	8	3.2%	52	7.1%	60	6.2%
<1500	123	49.8%	338	46.4%	461	47.3%
1500-2500	116	47.0%	338	46.4%	454	46.6%
>2500						
Apgar score at 1st minute	17	12.3%	48	12.8%	65	12.7%
0-6	103	74.6%	290	77.5%	393	76.8%
7-9	18	13.0%	36	9.6%	54	10.5%
10						
Apgar score at 5th minute	1	0.8%	14	3.9%	15	3.1%
0-6	41	31.1%	130	36.3%	171	34.9%
7-9	90	68.2%	214	59.8%	304	62.0%
10						
Breastfeeding duration	265	100.0%	0		265	26.3%
no			75	10.1%	75	7.5%
less than 2 weeks			189	25.5%	189	18.8%
2 to 6 weeks			181	24.4%	181	18.0%
6 weeks to 3 months			148	20.0%	148	14.7%
3 to 6 months			148	20.0%	148	14.7%
more than 6 months						

Note: Descriptive statistics of children included in the study. MCMA=monochorionic monoamniotic, MCDA=monochorionic diamniotic, DCDA=dichorionic diamniotic, SD=standard deviation, BMI=body mass index, IVF=in vitro fertilization. ICSI=intracytoplasmic sperm injection, SES=socio-economic status.

split up the sample into two groups (median age split) to have two equally-sized groups: children younger than 10 years, and children older than 10 years. Epigenome-wide significance was assessed following Bonferroni correction for the number of methylation sites tested (α=0.05/787,711=6.34x10^{-08}).

We carefully checked if epigenome-wide significant associations were influenced by outliers. As a sensitivity analysis, we repeated the association analysis in the discovery sample for significant CpGs without outliers defined using the Tukey method[88], in which an outlier is any value greater than the upper quartile plus three-times the interquartile range or lower than the lower quartile minus three-times the interquartile ranges.

Twin correlations

The correlation between DNA methylation levels of MZ and DZ twins was computed for CpGs that were significant in the discovery study, and which were robust to outliers. The

methylation beta-values were adjusted for the set of covariates as included in the basic EWAS model with a linear model and residuals were saved. Next, the MZ and DZ twin correlations were computed on the residuals.

Replication

CpGs that were significant in the discovery study, and which were robust to outliers, were selected for replication. The basic and adjusted models were applied as in the discovery sample: in NTR the GEE model had the same covariates; in ALSPAC, cell count estimates for peripheral blood were included instead of epithelium cells count; the sex, age of the child at blood collection around age 7 and the same covariates from the adjusted model 2 were included; no siblings were included. In ALSPAC, DNA methylation variation due to technical artefacts or unknown confounders were handled by including 20 surrogate variables generated from the DNA methylation data using the 'sva' R package[89] and associations were tested with linear models implemented by the limma package[90].

Methylation data annotation

To examine previously reported associations for epigenome-wide significant CpGs associated with breastfeeding, we looked them up in the EWAS atlas (https://bigd.big.ac.cn/ewas/tools; access October 3 2019[91]) and EWAS catalog (http://www.ewascatalog.org; access October 3 2019). To analyse the possible function of CpGs, we searched for overlaps with mQTLs in a study that analysed EPIC array data from 102 buccal samples[63], to identify the associated SNPs and looked up these SNPs and genes in the GWAS catalog (https://www.ebi.ac.uk/gwas/; access October 4 2019).

Overlap with previous findings

The follow-up analyses were done for CpGs previously associated with breastfeeding in EWAS or candidate gene studies. It included all CpGs on the EPIC array that are annotated to the genes *LEP* (nCpGs=23)[52–55,58], *IL4R* (nCpGs=37)[58], *CDKN2A* (nCpGs=32)[59], 1 CpG from the study of Hartwig et al.[92], and 4,297 suggestive CpGs from the study of Naumova et al.[53]. In total, this resulted in a list of 4,370 CpGs, of which 3,859 CpGs were present (after QC) in our data. For these CpGs, we performed a look-up analysis in the total sample (n =1,006) and in the subsample younger than10 years (n=517) of the discovery NTR cohort. Significance was assessed after Bonferroni correction for the number of CpGs tested (α=0.05/3,858=1.29x10^{-05}).

Results

Descriptive statistics

Descriptive statistics are presented in **Table 1**. The total discovery sample consisted of 1,006 children (mean age 9.5, SD=1.89); 73.7% of the children were breastfed with different duration, and 26.3% never received breastfeeding. Most of the children were breastfed at least one month: 189 (25.5%) were breastfed for 2 weeks to 1.5 months, 181 (24.4%) were breastfed for 1.5-3 months, 148 (20%) were breastfed for 2-6 months,

and 148 (20%) were breastfed more than 6 months. The subsample of children younger than 10 years old included 517 twins (mean age 7.9, SD=1.1) (see **Supplemental Table S4**) and the group older than 10 years included 489 twins (mean age=11.16, SD=0.72) (see **Supplemental Table S5**). The majority of twin pairs was concordant for breastfeeding (99.4%). Among the discordant twin pairs were 3 pairs of which one twin was not breastfed while co-twin received breastfeeding, and 3 pairs of which the co-twins experienced different durations of breastfeeding.

Breastfeeding duration (6 categories) was significantly positively associated with socio-economic status (β=0.353, SE=0.09, *p*=0.0004) and inversely associated with maternal smoking (β=-0.733, SE=0.32, *p*=0.02) and gestational age (β=-0.194, SE=0.09, *p*=0.03) (see **Supplemental Table S3**). Breastfeeding was not significantly correlated with maternal pre-pregnancy BMI (β=-0.117, SE=0.75, *p*=0.119), maternal age at delivery (β=-0.016, SE=0.09 *p*=0.86), and cell composition of buccal swabs (β=1.27, SE=1.2, *p*=0.25 for count of epithelial cells, and β=14.78, SE=9.8, *p*=0.13 for count of natural killer cells).

Association analysis findings

First, genome-wide DNA-methylation analyses were performed to test the association between breastfeeding (never/ever) and DNA-methylation level, while adjusting for sex, age at DNA sample collection, estimated proportion of epithelial cells, estimated proportion of natural killer cells, row of the sample on the chip and bisulfite plate (model 1). No epigenome-wide significant sites were identified (see **Supplemental Table S8**). Genome-wide EWAS test statistics showed no inflation (see **Supplemental Figure S1**). DNA methylation was also not associated with duration of breastfeeding (see **Supplemental Table S9**).

In the EWAS in children younger than 10 years with the same basic model, we identified four epigenome-wide significant CpGs (see **Supplemental Table S10**): cg25178826 (β=-0.026, SE=0.003, $p=8.04\times10^{-12}$) is located in the 5'UTR region of *PRLR* (prolactin receptor) gene, cg12087956 (β=-0.03, SE=0.005, $p=1.8\times10^{-08}$) in the gene body of *CDAN1* (codanin 1), cg24192772 (β=-0.024, SE=0.004, $p=2.5\times10^{-08}$) in the gene body of *FOXK2* (forkhead box K2), and cg10142656 (β=-0.02, SE=0.004, $p=6.28\times10^{-08}$) is mapped to the transcription start site (TSS1500) of *TRMT10B* (tRNA methyltransferase 10B). These four CpGs were not strongly associated with breastfeeding duration (see **Supplemental Table S11**). Plots of the methylation values for these 4 CpGs revealed several extreme methylation values in the data (see **Supplemental Figure S2**, **Supplemental Table S10**), and the association with breastfeeding disappeared after outlier removal (cg25178826 β=-0.00006, SE=0.001, *p*=0.98; cg12087956 β=.00004, SE=0.0009, *p*=0.5; cg24192772 β=0.0009, SE=0.001, *p*=0.15; cg10142656 β=.0001, SE=0.0008, *p*=0.13).

Next, we performed EWASs with adjustment for SES, maternal age at delivery, maternal pre-pregnancy BMI, maternal prenatal smoking, and gestational age (model 2). One CpG cg22491379 was significant in the total sample (β=-0.007, SE=0.001, $p=1.3\times10^{-09}$) (see **Figure 2a**, **Table 2**, **Supplemental Table S12**), and seven CpGs were significant in the children younger than 10 years (see **Table 2**, **Supplementary Table S14**). The Manhattan plot for the EWAS in children younger than 10 years showed a peak on chromosome 21,

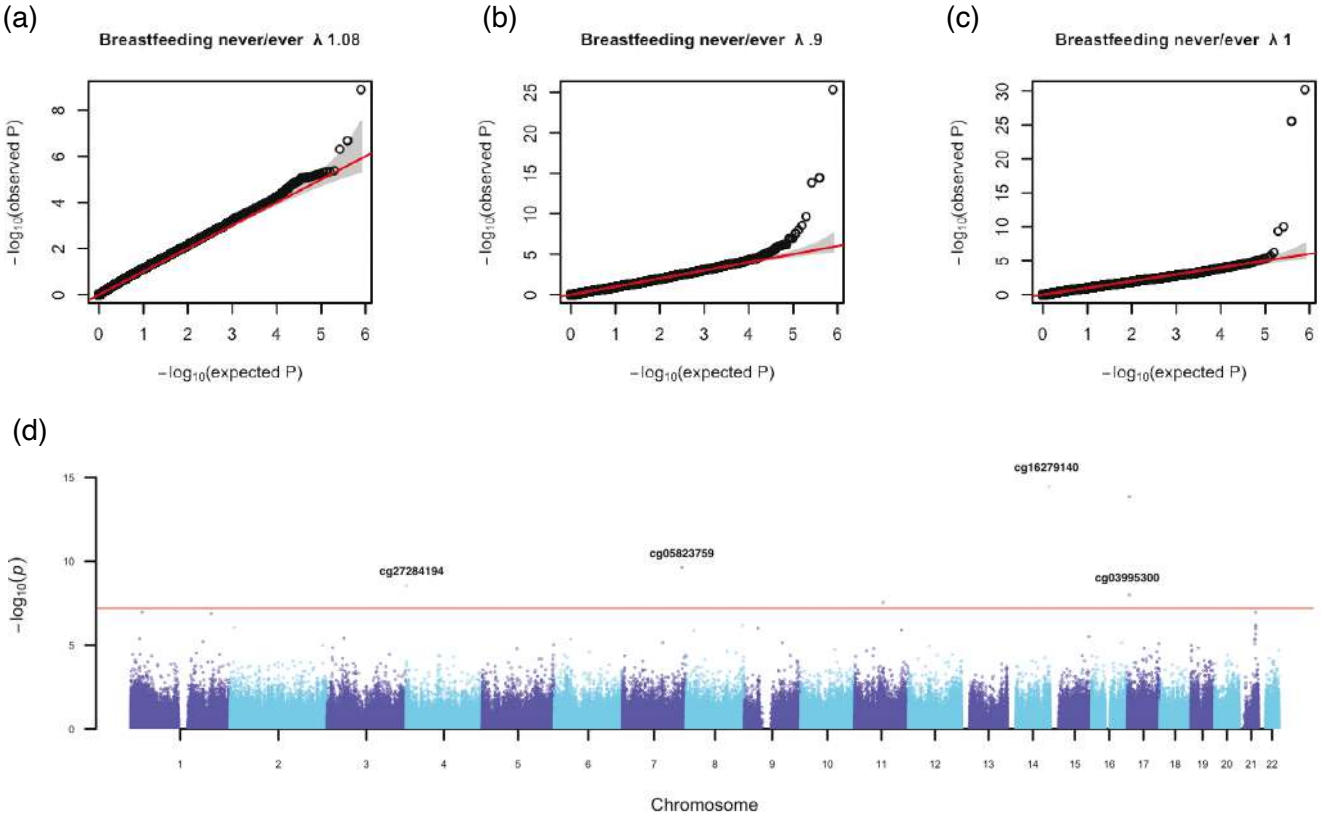


Figure 2. QQ plots: **(a)** in the total sample; **(b)** in the subsample of children younger than 10 years old; **(c)** in the subsample of children older than 10 years old; **(d)** Manhattan plot shows the epigenome-wide association study results of breastfeeding 'never'/'ever' in children younger than 10 years old. Covariates included sex, age, SES, maternal age at delivery, maternal pre-pregnancy BMI, maternal prenatal smoking, gestational age, EPIC array row and bisulfite sample plate, and cell composition. In the Manhattan plot, the red line represents the Bonferroni threshold ($6.34x10^{-08}$). Given CpG names indicate significant loci after removal of outliers.

TABLE 2.
Summary of epigenome-wide significant CpGs from the discovery EWAS of DNA signatures of breastfeeding

cgID	chromo-some	position	gene	gene region	Discovery study			Discovery study without outliers		
					Estimate	SE	*P* value	Estimate	SE	*P* value
Basic model (1). Sub-sample <10 years (*n*=517)										
cg25178826	chr5	35165447	*PRLR*	5'UTR	-0.026	0.004	$8.04x10^{-12}$	$-6.04x10^{-05}$	0.001	0.98
cg12087956	chr15	43022167	*CDAN1*	Body	-0.031	0.005	$1.18x10^{-08}$	$4.92x10^{-05}$	0.001	0.50
cg24192772	chr17	80536920	*FOXK2*	Body	-0.024	0.004	$2.52x10^{-08}$	$9.17x10^{-04}$	0.001	0.15
cg10142656	chr9	37753047	*TRMT10B*	TSS1500	-0.019	0.004	$6.28x10^{-08}$	$9.98x10^{-05}$	0,0007	0.14
Adjusted model (2). Discovery sample (n = 1,006)										
cg22491379	chr2	120553625	*PTPN4*	5'UTR	-0.007	0.001	$1.30x10^{-09}$	-0.005	0.001	$5.78x10^{-03}$
Adjusted model (2). Sub-sample <10 years (*n*=517)										
cg03463465	chr6	164143581			0.360	0.034	$4.51x10^{-26}$	-0.003	0.001	0.01
cg07670516	chr17	5019840	*ZNF232*	5'UTR	0.249	0.032	$1.40x10^{-14}$	0.006	0.014	0.65
cg20820810	chr11	71850130	*FOLR3*	Body	-0.300	0.054	$2.82x10^{-08}$	-0.001	0.001	0.21
cg16279140	chr14	103981749			-0.411	0.052	$3.50x10^{-15}$		no outliers	
cg05823759	chr7	149646627			0.205	0.032	$2.35x10^{-10}$		no outliers	
cg27284194	chr4	1044797			0.638	0.107	$2.90x10^{-09}$		no outliers	
cg03995300	chr17	5019989	*ZNF232*	5'UTR	0.229	0.040	$1.02x10^{-08}$		no outliers	

Note: α= 0.05/787,711=6.34e-08. Basic model 1 included breastfeeding coded as 'never' and 'ever' with adjustment for sex, age at DNA methylation, count of epithelial cells, count of natural killer cells, EPIC array row and bisulfite sample plate. Adjusted model (2) included in additional to basic model covariates: SES, maternal smoking during pregnancy, maternal age at birth, maternal pre-pregnancy BMI, gesta-tional age. In bold: CpGs selected for replication.

which contains a cluster of CpGs just below the epigenome-wide significance threshold **(see Figure 2d**).

Genome-wide EWAS test statistics from the analyses of model 2 (in the total sample and in the children younger than 10 years) showed no inflation (see **Figure 2a, 2b**). For four out of the eight significant CpGs, the association disappeared after removal of outliers (see **Supplemental Figures S4-5**). The four CpGs that were unaffected by outliers were selected for replication analysis (see **Supplemental Figure S6**): cg16279140 (β=-0.4, SE=0.05, p=3.5x10^{-15}), cg05823759 (β=0.2, SE=0.032, p=2.35x10^{-10}), cg27284194(β=0.64, SE=0.12, p=2.9x10^{-09}), cg03995300 (β=0.23, SE=0.04, p=1.2x10^{-08}) (see **Supplemental Table 14**). All four CpGs were significant in children younger than 10 years.

We next examined the association with breastfeeding in children older than 10 (model 2). In this analysis, genome-wide test statistics showed mild inflation (λ=1.11). After adjusting for inflation, four CpGs remained epigenome-wide significant, but none were significant after outlier removal. The four CpGs significant in the children younger than 10 years, had smaller effects and 3 CpGs showed inverse directions of effect in children older than 10: cg16279140 (β=.04, SE=0.02, p=0.04), cg05823759 (β=-.002, SE=.02 p=.92), cg27284194 (β=0.18, SE=0.06, p=0.002), and cg03995300 (β=-0.003, SE=0.02, p=0.89) (see **Supplemental Table S16**). The quantile-quantile plots (see **Figure 2b** and **2c**) suggested a stronger association signal for breastfeeding in children younger than 10 years than in an older group. We computed the correlations between the regression coefficients (i.e. the methylation difference associated with breastfeeding) for the top 100 CpGs from each age group. These correlations were weak (r=0.255, p=0.10 for the top 100 CpG from the EWAS in children < 10 years and from the EWAS in children > 10 years). This suggests that methylation profiles associated with breastfeeding are different in children younger than 10 years and older than 10 years.

We computed twin correlations in MZ and DZ pairs for the significant CpGs, as we assume that equal exposure to breastfeeding should cause similarity in methylation level of twins. Correlations of methylation levels of all four sites were very high in MZ twins, and almost twice as high as in DZ twins: cg16279140 r_{MZ}=.906, r_{DZ}=.505, cg05823759 r_{MZ}=.953, r_{DZ}=.609, cg27284194 r_{MZ}=.972, r_{DZ}=.462, and cg03995300 r_{MZ} = .902, r_{DZ}=.453. This pattern suggests that there are heritable influences on DNA methylation at these CpGs, as previously shown by other studies of DNA methylation heritability.

Replication analysis of our findings in other samples

Both replication datasets were comparable to the discovery subsample of children younger than 10 years old (see **Supplementary Tables S4, S6, S7**). As expected, the singleton children (ALSPAC) had higher gestational age (~39 weeks vs ~35.7), higher birthweight (~3500g vs ~2400g) and lower proportion of caesarean deliveries (~9% vs ~36.8%). We performed replication analysis for significant CpGs unaffected by outliers from the adjusted model 2 of the discovery analysis with the same set of covariates (see **Table 3**).

In NTR replication cohort, for one CpG, the direction of effect changed (cg03995300, ZNF232, β=-0.050, SE=.033, p=0.013), three other CpGs had the same direction of effect

TABLE 3.
Summary of association between breastfeeding and significant in discovery study CpGs in repli-cation

cgID	Direction of effect in discovery study <10 years	NTR replication study (*n*=98)			ALSPAC replication study (*n*=938)		
		Estimate	SE	*P* value[a]	Estimate	SE	*P* value
cg16279140	-	-0.0326	0.0412	0.43	NA	NA	NA
cg05823759	+	0.0329	0.0332	0.32	NA	NA	NA
cg27284194	+	0.0668	0.0542	0.21	0.0047	0.016	0.77
cg03995300	+	-0.0502	0.0334	0.13	0.0140	0.011	0.19

Note: Adjusted model (2) is used. α=0.05/4=0.01. NA=not available on 450k platform.

but were not significant: cg16279140 (β=-0.03, SE=0.04, *p*=0.43), cg05823759 (β=0.03, SE=.033, *p*=0.32), cg27284194 (β=0.07, SE=0.054, *p*=0.21) (see **Table 3**).

In ALSPAC, the most important difference to note were the different tissues used for DNA methylation profiling (peripheral blood) and the platform (the Illumina Infinium HumanMethylation450 Beadchip). Because of the platform difference, roughly half of the CpG sites included in the discovery DNA methylation profiles were measured in ALSPAC. In the adjusted model 2, two of the four most strongly associated CpG sites were not included on the ALSPAC platform (cg16279140 and cg05823759) nor were any other sites within 1000bp of these sites. The other two sites, cg27284194 and cg03995300, were included, but neither association was replicated ($p > 0.19$ and $p > 0.77$, respectively) in the adjusted model (see **Table 3**). An overall lack of replication was confirmed by low correlations between effect estimates in the discovery and ALSPAC. In particular, we identified the 100 most strongly associated CpG sites in the discovery that were also present in the ALSPAC DNA methylation profiles. Correlation of effect estimates for these sites between studies was low (Pearson's *R*=0.13 with *p*=0.2 for the adjusted model).

Replication analysis of findings from previous EWAS

Of the 3,858 CpGs from previous literature (see **Supplemental Tables S17**, **S18**), four CpGs were associated with breastfeeding in our sample ($x=1.29x10^{-05}$, see **Table 4**). One site was significant in the total sample: cg16387046 located on chromosome 12 in the *MUCL1* (mucin like 1) gene (β=0.027, SE=0.005, $p=4.9x10^{-07}$). Three sites were significant in the children younger than 10 years: cg16704958 (β=0.009, SE=0.002, $p=8.03x10^{-06}$) and cg11287055 (β=0.056, SE=0.01, $p=4.9x10^{-06}$) located on chromosome 21 in the VPS26C (endosomal protein sorting factor C; previous name *DSCR3*) gene, and cg26479305 (β =0.338, SE=0.07, $p=1.11x10^{-05}$) located on chromosome 12 in the *ATG101* (autophagy related 101, previous name *C12orf44*) gene. All these CpGs were previously reported in the association study by Naumova et al.[53] carried out in infants around 2 years old (mean age 25.7 months) with DNA methylation profiling in peripheral

TABLE 4.
Replication of CpGs from previous literature

cgID	chromosome	position	gene	gene region	Estimate	SE	
			Total sample (n=1,006)				***P* value**
cg16387046	chr12	55248207	*MUCL1*	TSS200	0.027	0.005	$4.93x10^{-07}$
			Sub-sample <10 (n=517)				
cg11287055	chr21	38630234	*VPS26C (DSCR3)*	Body	0.056	0.012	$4.93x10^{-06}$
cg16704958	chr21	38630728	*VPS26C (DSCR3)*	Body	0.009	0.002	$8.03x10^{-06}$
cg26479305	chr12	52470979	*ATG10 (C12orf44)*	3'UTR	0.338	0.077	$1.11x10^{-05}$

Note: α=0.05/3859 =$1.29x10^{-05}$. The table shows all CpGs that were previously reported to be associated with breastfeeding and that were significantly replicated by our study.

blood with Illumina EPIC BeadChip. The direction of association was positive for all four CpGs in both studies. CpGs located in/nearby *LEP*[52,55], *IL4R*[58] and *CDKN2A*[59], previously discussed as candidate genes for association with breastfeeding, were not significant in our study.

Methylation data annotation

For the significant CpGs from discovery and follow-up studies, we looked up the associations with nearby genetic variants (mQTLs) in the results from a previously published mQTL study of buccal-derived DNA in monozygotic twins[63]. Four CpGs were associated with mQTLs (cg27284194, associated with SNPs on chromosome 4: 927973-1039876; cg16279140, associated with SNPs on chromosome 14: 103823243-104186876; cg16387046, associated with SNPs on chromosome 12: 54830518-54896008; cg26479305, associated with SNPs on chromosome: 12: 52492131), and for four CpGs (cg03995300, cg05823759, cg11287055, cg16704958) no significant mQTLs were found.

Next, we compared our results against all previously associated traits in EWASs (EWAS atlas and EWAS catalog) and GWASs (GWAS catalog). One CpG (cg03995300) is mapped to *ZNF232* and was previously associated with prenatal maternal smoking, sex, and ancestry[56,93,94]. The *ZNF232* locus has been associated with family history of Alzheimer's disease in a GWA meta-analysis[95]. The intergenic CpG on chromosome 4 (cg27284194) has been previously associated with infertility[96,97]. In GWAS, the region 4:927973-1039876 (harbouring mQTLs for cg27284194) has been associated with a number of traits and diseases, including bone mineral density[98–102] and several metabolomic characteristics such as blood protein[103,104] and triglyceride[105] levels (See **Supplemental Table S19**). The intergenic CpG on chromosome 14 (cg16279140) did not appear in previous EWASs. In GWASs, SNPs on chromosome 14: 103823243-104186876 have been linked to amino acids levels as biomarkers of metabolic disorders[106] and other blood metabolites[107], and some disease and addictions including multisite chronic pain[108], metabolite changes in chronic kidney disease[109], bipolar disorder[110], alcohol consumption[111] and risk-taking

behavior[110] (See **Supplemental Table S19**). The intergenic CpG on chromosome 7, cg05823759, has not been previously identified in EWAS and was not associated with mQTLs.

Four CpGs that were replicated from previous studies on breastfeeding exposure, have been reported in association with other traits and diseases. The *ATG10 (C12orf44)* gene, where cg26479305 is located, has been related to autophagy pathway and cellular senescence in EWAS[112]. It was related to regulators of inflammation (circulating cytokines and growth factors)[113], hematological traits in GWASs[114], and prenatal arsenic exposure[114].

CpGs (cg11287055, cg16704958) mapped to the *VPS26C (DSCR3)* have not been reported in other EWASs. *VPS26C* is a component of the retriever complex, which plays role in cell surface processes such as cell migration, cell adhesion, nutrient supply and cell signaling[115].

Discussion

We aimed to study DNA methylation in buccal cells in mid-childhood in association with breastfeeding as an exposure. Associations were tested in the total sample and in groups stratified by age. The age distribution allowed us to analyse if effects of breastfeeding can be attenuated with age as this has been observed for other exposures in peripheral blood, including maternal smoking[56], birthweight[57] and breastfeeding[55].

We did not find evidence for robust epigenome-wide significant associations in the total sample (1,006 children, age 5-12) but did observe associations in the younger age group (before 10 years) that did not appear in the group of children of 10-12 years old. This suggests that epigenetic alterations in certain genomic regions might be associated with nutritional differences in the early postnatal life that tend to fade across childhood. However, since the CpGs were not replicated in the small buccal-cells replication NTR dataset and in the large blood-cells replication dataset from ALSPAC, further studies are required to follow up on these findings. It also remains to be examined if the associations reflect a causal effect of breastfeeding on methylation, and whether these methylation differences might influence developmental trajectories and later-life health of the child.

The observation that associations between breastfeeding and methylation at some CpGs are age-dependent may be explained by age-related changes in DNA methylation that have been reported in several studies[116,117], including children in the age range of our study[118–122]. Similarly, Sherwood et al.[55] observed that breastfeeding is associated with DNA methylation of the *LEP* gene at age 10 but not at age 18. Previous studies have identified DNA methylation signatures of prenatal nutrition in peripheral tissues at various life stages, even in adulthood. DNA methylation signatures in many genomic regions were identified in middle-aged individuals who were exposed to the 1944/45 Dutch hunger winter at the time of conception and many of these sites were related to growth, developmental processes and metabolism[18,19]. It is unclear if effects of early life postnatal exposures can have similar long-term effects as prenatal exposures in humans, as most

studies of early life influences on the epigenome have focused on prenatal exposures, or have not examined long-term effects. None of the genes reported in studies of prenatal malnutrition *(SMAD7, CDH23, INSR, RFTN1, CPT1A, KLF13*[19]*, IGF2*[18]*, LEP*[123]*)* were present among the top CpGs associated with breastfeeding in our study. Several explanations can be proposed. First, postnatal nutrition and prenatal nutrition exposures might influence the DNA methylation differently. Second, DNA methylation signatures induced by early nutrition exposure can be different across tissues examined (buccal cells in our study versus peripheral blood in previous famine studies). Third, the prenatal famine study examined effects of an extreme exposure.

The presence of associations with breastfeeding initiation (breastfeeding never vs ever), and the absence of findings with breastfeeding duration suggests that the association depends on the exposure to breast milk rather than its duration. This could potentially be in line with previous findings showing that effects of exposures on DNA methylation occur only when the individual is exposed in certain sensitive life periods[118]. Our finding of stronger epigenetic associations with breastfeeding ever vs. never than with duration of breastfeeding is also in line with studies providing evidence that any breastfeeding has stronger biological effects than the duration of breastfeeding, for instance the impact of first maternal milk (colostrum) on immunoglobulins and further neonatal health, especially for small for gestational age and low-birthweight infants[22,24].

We observed associations at a total of eight CpG sites: cg16279140, chromosome 14, intergenic; cg05823759, chr7, intergenic; cg27284194, chromosome 4, intergenic; cg03995300, chromosome 17, *ZNF232;* cg16387046, chromosome 12, *MUCL1;* cg11287055, chromosome 21, *VPS26C (DSCR3);* cg16704958, chromosome 21, *VPS26C (DSCR3);* cg26479305, chromosome 12, *ATG10 (C12orf44).* Methylation levels at these CpGs are potentially influenced by both environmental influences, such as breastfeeding, and by genetic variation, although an alternative interpretation could be that the genomic regions of these CpGs affect breastfeeding function. Unfortunately, no GWAS on breastfeeding is available to verify this suggestion. It remains to be examined whether methylation differences induced by breastfeeding have an effect on the aforementioned traits. Interestingly, associations with breastfeeding were reported previously in epidemiological studies for bone mineral content and bone mineral density[124] and metabolite profiles[125]; outcomes that have been associated with SNPs in the regions detected in our study as differentially methylated between breastfed and non-breastfed children. It remains to be examined whether methylation differences induced by breastfeeding have an effect on these traits.

Since the role of nutrition is systemic, it is assumed that biomarkers must be present in different tissues. To study epigenetic mechanisms of breastfeeding, previous studies have examined DNA from peripheral blood[52–55,58] and also from tumor tissue[59]. Buccal epithelium is of interest because it offers a non-invasive way of biosample collection for epigenetic analysis. A number of studies have demonstrated the potential of buccal cells to study DNA methylation[126]. We previously showed that although there is some correlation between DNA methylation in buccal and blood cells (n=22, age=18 years), it is low for most CpGs interrogated by the Illumina 450k array[127]. The methylation levels of two of the CpG sites that were included in the ALSPAC replication study are highly

correlated in buccal and blood cells: cg27284194 (r=.864) and cg03995300 (r=.782). In spite of this, the associations with breastfeeding observed in NTR were not replicated in ALSPAC. We observed some overlap with the study of Naumova et al.[53] in peripheral blood that used the EPIC array, and these sites had medium correlations between DNA methylation in buccal and blood cells (cg16387046 r=0.690, cg11287055 r=0.413, cg16704958 r=0.328).

We observed that breastfeeding-associated CpGs show strong correlations in MZ twins, and larger correlations in MZ twins compared to DZ twins. As the large majority of twin pairs in our study were almost always concordant for breastfeeding, it is expected that twins of both types show resemblance for DNA methylation levels at these sites. The larger correlation in MZ compared to DZ twins in our study suggests that these sites are also subject to heritable influences. In line with this observation, some of these CpGs are associated with mQTLs. It should be noted that breastfeeding and lactation itself are heritable traits. In previous twin studies the heritability of initiation of breastfeeding ranged from 49%[128] to 70%[129]. DNA-methylation profiles may be influenced by several early life factors. In our study, almost all early environmental factors of interest were shared by twins: SES, maternal age at birth, maternal pre-pregnancy BMI, maternal smoking during pregnancy and gestational age.

Some CpGs that did not reach epigenome-wide significance can have potential for further epigenetic studies of breastfeeding. The CpG cg22491379, located in the *PTPN4* on chromosome 2, was discovered in total sample of children of 5–12 years old and after outlier removal had suggestive significance (p=5.78x10^{-03}). *PTPN4* was listed in suggestive regions associated with breast morphology (breast size)[130]. Some top CpGs in the discovery study (non-adjusted model) are associated with genes that are involved in growth and metabolism. The *PRLR* (prolactin receptor) gene, located on chromosome 5, is involved in prolactin receptor activity and growth hormone receptor signalling pathway. Growth hormone binds to the prolactin receptor, this being the basis of induction of lactation by growth hormone[131]. The *FOXK2* gene, located on chromosome 17, is a member of forkhead box transcription factors and is involved in glucose metabolism, aerobic glycolysis and autophagy[132], playing a role in metabolic reprogramming towards aerobic glycolysis[133]. It was associated lean body mass[134] and found to be hypomethylated in CpG islands in obese patients' adipose tissues[135]. We observed a cluster of associations on chromosome 21, each individual association just below epigenome-wide significance. The *VPS26C (DSCR3)* gene is located on chromosome 21 and, as mentioned earlier, contains two CpGs previously associated with breastfeeding in blood[53].

The strengths of our study include the use of the buccal epithelium methylome to investigate breastfeeding for the first time, sample size and the availability of replication data, including data from another tissue (blood). Furthermore, our exposure, breastfeeding, was assessed very shortly after it occurred, reducing the risk of measurement error and recall bias. Both the discovery and replication samples have been thoroughly phenotyped and studied extensively. Our study also has limitations. First, the findings in the discovery study were not replicated in the cohort of children with buccal cell DNA methylation of the same age, possibly due to the small size of this cohort (98 monozygotic twins). Second, the discovery sample was included in a study of aggressive behavior, and not

primarily examined for the purposes of research on breastfeeding. Third, we did not have longitudinal DNA methylation data to measure the stability of effects; however, we examined possible attenuation of breastfeeding effects through age-stratified analysis. Fourth, there are currently no datasets available on gene expression in buccal cells, thus we could not examine relationships between DNA methylation and transcription. A strength of the current analysis is that it used the Illumina EPIC array, which has much greater coverage than the 450k array. Importantly, some of the top CpGs identified in the discovery EWAS were all novel EPIC probes. Replication analysis in ALSPAC on DNA methylation data from peripheral blood, however, used the 450k array, and therefore did not permit look-up of the same CpGs. The findings of this study require further replication in cohorts with buccal epithelium and other tissues to improve our understanding of breastfeeding-associated methylation changes in different tissues, and the possible utility as a biomarker of early life nutrition. Finally, a difficulty in studies of breastfeeding is that breast milk composition is unique in each mother. Lactation is influenced not only by genetic variation, but also by many environmental factors such as the mother's nutrition, lifestyle, level of stress, and attachment to the child[24,25,136–139]. Future epigenetic studies of breastfeeding could stratify the breastfeeding sample on the basis of criteria of breast milk composition. This might be more informative for predicting a child's outcome, but also be a better indicator of a mother's well-being. The value of our research is that it combined breastfeeding, the prenatal characteristics and methylation data. In the future, results from studies that will integrate epigenomic data with genomics, transcriptomics, and metabolomics may be used to develop prediction models of long-term outcomes in child development and health. Understanding of the mechanisms associated with breastfeeding will help to develop interventions to improve children's health and reduce risk of chronic disease by supporting breastfeeding or optimizing infant nutrition when breastfeeding is not possible.

Conclusions

Our study provided a first indication that breastfeeding as an early life environmental factor may be associated with epigenetic variation in buccal cells in children. The findings point at new candidate loci influenced by breastfeeding. Future studies are needed to investigate if the DNA methylation signatures are caused by breastfeeding or by other unmeasured confounders (including genetic predisposition to give or receive breastfeeding, other aspects of prenatal or postnatal diet), whether they are influenced by percentage of breastfed meals, exclusive breastfeeding duration, breastmilk composition etc. and what age-related mechanisms drive changes in the association between breastfeeding and methylation.

References

1. Godfrey, K. M., Lillycrop, K. A., Burdge, G. C., Gluckman, P. D. & Hanson, M. A. Epigenetic Mechanisms and the Mismatch Concept of the Developmental Origins of Health and Disease. *Pediatric Research* **61**, 5R-10R (2007).
2. Gluckman, P. D., Hanson, M. A. & Low, F. M. The role of developmental plasticity and epigenetics in human health. *Birth Defects Research Part C: Embryo Today: Reviews* **93**, 12–18 (2011).
3. Verduci, E. *et al.* Epigenetic Effects of Human Breast Milk. *Nutrients* **6**, 1711–1724 (2014).
4. Langley-Evans, S. C. Nutrition in early life and the programming of adult disease: a review. *Journal of Human Nutrition and Dietetics* **28**, 1–14 (2015).
5. Hochberg, Z. *et al.* Child health, developmental plasticity, and epigenetic programming. *Endocrine reviews* **32**, 159–224 (2011).
6. Raghuraman, S., Donkin, I., Versteyhe, S., Barrès, R. & Simar, D. The Emerging Role of Epigenetics in Inflammation and Immunometabolism. *Trends in Endocrinology & Metabolism* **27**, 782–795 (2016).
7. Paparo, L. *et al.* The influence of early life nutrition on epigenetic regulatory mechanisms of the immune system. *Nutrients* **6**, 4706–19 (2014).
8. Canani, R. B. *et al.* Epigenetic mechanisms elicited by nutrition in early life. *Nutrition research reviews* **24**, 198–205 (2011).
9. Wadhwa, P., Buss, C., Entringer, S. & Swanson, J. Developmental Origins of Health and Disease: Brief History of the Approach and Current Focus on Epigenetic Mechanisms. *Seminars in Reproductive Medicine* **27**, 358–368 (2009).
10. Gluckman, P. D., Hanson, M. A. & Beedle, A. S. Early life events and their consequences for later disease: A life history and evolutionary perspective. *American Journal of Human Biology* **19**, 1–19 (2007).
11. Bateson, P. et al. Developmental plasticity and human health. *Nature* **430**, 419–421 (2004).
12. Choi, S.-W. & Friso, S. Epigenetics: A New Bridge between Nutrition and Health. *Advances in Nutrition* **1**, 8–16 (2010).
13. Peter, C. J. et al. DNA Methylation Signatures of Early Childhood Malnutrition Associated With Impairments in Attention and Cognition. *Biological psychiatry* **80**, 765–774 (2016).
14. Waterland, R. A. & Michels, K. B. Epigenetic Epidemiology of the Developmental Origins Hypothesis. *Annual Review of Nutrition* **27**, 363–388 (2007).
15. Cutfield, W. S., Hofman, P. L., Mitchell, M. & Morison, I. M. Could Epigenetics Play a Role in the Developmental Origins of Health and Disease? *Pediatric Research* **61**, 68R-75R (2007).
16. Tammen, S. A., Friso, S. & Choi, S.-W. Epigenetics: The link between nature and nurture. *Molecular Aspects of Medicine* **34**, 753–764 (2013).
17. Faa, G. *et al.* Fetal programming of neuropsychiatric disorders. *Birth Defects Research Part C: Embryo Today: Reviews* **108**, 207–223 (2016).
18. Heijmans, B. T. *et al.* Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17046–9 (2008).
19. Tobi, E. W. *et al.* DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nature Communications* **5**, 5592 (2014).
20. Barker, D. J. The fetal and infant origins of adult disease. *BMJ* **301**, 1111–1111 (1990).
21. Ruiz, L., García-Carral, C. & Rodriguez, J. M. Unfolding the Human Milk Microbiome Landscape in the Omics Era. *Frontiers in Microbiology* **10**, 1378 (2019).
22. Fanos, V., Reali, A., Marcialis, M. A. & Bardanzellu, F. What you have to know about Human Milk Oligosaccharides. *Journal of Pediatric and Neonatal Individualized Medicine (JPNIM)* **7**, e070137 (2018).

23. Hassiotou, F. *et al.* Breastmilk Is a Novel Source of Stem Cells with Multilineage Differentiation Potential. *STEM CELLS* **30**, 2164–2174 (2012).
24. Bardanzellu, F., Fanos, V. & Reali, A. "Omics" in Human Colostrum and Mature Milk: Looking to Old Data with New Eyes. *Nutrients* **9**, (2017).
25. Dessì, A. *et al.* Metabolomics of Breast Milk: The Importance of Phenotypes. *Metabolites* **8**, 79 (2018).
26. Forsum, E. & Lonnerdal, B. Variation in the contents of nutrients of breast milk during one feeding. *Nutrition reports international* **19**, 815–20 (1979).
27. Victora, C. G. *et al.* Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *The Lancet* **387**, 475–490 (2016).
28. McInerny, T. K. Breastfeeding, Early Brain Development, and Epigenetics—Getting Children Off to Their Best Start. *Breastfeeding Medicine* **9**, 333–334 (2014).
29. Ip, S. *et al.* Breastfeeding and maternal and infant health outcomes in developed countries. *Evidence report/technology assessment* 1–186 (2007).
30. Wisnieski, L., Kerver, J., Holzman, C., Todem, D. & Margerison-Zilko, C. Breastfeeding and Risk of Metabolic Syndrome in Children and Adolescents: A Systematic Review. *Journal of Human Lactation* **34**, 515–525 (2018).
31. Chirico, G., Marzollo, R., Cortinovis, S., Fonte, C. & Gasparoni, A. Antiinfective Properties of Human Milk. *The Journal of Nutrition* **138**, 1801S-1806S (2008).
32. Martin, C., Ling, P.-R. & Blackburn, G. Review of Infant Feeding: Key Features of Breast Milk and Infant Formula. *Nutrients* **8**, 279 (2016).
33. Else-Quest, N. M., Hyde, J. S. & Clark, R. Breastfeeding, Bonding, and the Mother-Infant Relationship. *Merrill-Palmer Quarterly* vol. 49 495–517 (2003).
34. WHO | Protecting, promoting, and supporting breastfeeding in facilities providing maternity and newborn services: the revised Baby-friendly Hospital Initiative 2018. *WHO* (2019).
35. Michaelsen, K. F., Lauritzen, L. & Mortensen, E. L. Effects of Breast-feeding on Cognitive Function. in *Breast-Feeding: Early Influences on Later Health* vol. 639 199–215 (Springer Netherlands, 2009).
36. Kramer, M. S. *et al.* Breastfeeding and Child Cognitive Development. *Archives of General Psychiatry* **65**, 578 (2008).
37. Horta, B. L., Loret de Mola, C. & Victora, C. G. Breastfeeding and intelligence: a systematic review and meta-analysis. *Acta Paediatrica* **104**, 14–19 (2015).
38. Blesa, M. *et al.* Early breast milk exposure modifies brain connectivity in preterm infants. *NeuroImage* **184**, 431–439 (2019).
39. Bartels, M., van Beijsterveldt, C. E. M. & Boomsma, D. I. Breastfeeding, maternal education and cognitive function: a prospective study in twins. *Behavior genetics* **39**, 616–22 (2009).
40. Harder, T., Bergmann, R., Kallischnigg, G. & Plagemann, A. Duration of Breastfeeding and Risk of Overweight: A Meta-Analysis. *American Journal of Epidemiology* **162**, 397–403 (2005).
41. Burke, V. et al. Breastfeeding and Overweight: Longitudinal Analysis in an Australian Birth Cohort. *The Journal of Pediatrics* **147**, 56–61 (2005).
42. Gillman, M. W. et al. Risk of Overweight Among Adolescents Who Were Breastfed as Infants. *JAMA* **285**, 2461 (2001).
43. Owen, C. G., Whincup, P. H., Odoki, K., Gilg, J. A. & Cook, D. G. Infant Feeding and Blood Cholesterol: A Study in Adolescents and a Systematic Review. *PEDIATRICS* **110**, 597–608 (2002).
44. Shoji, H. & Shimizu, T. Effect of human breast milk on biological metabolism in infants. *Pediatrics International* **61**, 6–15 (2019).

45. Klement, E., Cohen, R. V, Boxman, J., Joseph, A. & Reif, S. Breastfeeding and risk of inflammatory bowel disease: a systematic review with meta-analysis. *The American Journal of Clinical Nutrition* **80**, 1342–1352 (2004).
46. Amitay, E. L. & Keinan-Boker, L. Breastfeeding and Childhood Leukemia Incidence. *JAMA Pediatrics* **169**, e151025 (2015).
47. Horta, B. L., Loret de Mola, C. & Victora, C. G. Long-term consequences of breastfeeding on cholesterol, obesity, systolic blood pressure and type 2 diabetes: a systematic review and meta-analysis. *Acta Paediatrica* **104**, 30–37 (2015).
48. Lodge, C. *et al.* Breastfeeding and asthma and allergies: a systematic review and meta-analysis. *Acta Paediatrica* **104**, 38–53 (2015).
49. Horta, B. L. & Victora, C. G. Short-Term E ects of Breastfeeding: *A Systematic Review on the Benefits of Breastfeeding on Diarrhoea and Pneumonia Mortality. World Health Organization* (World Health Organization, 2013).
50. Fall, C. H. *et al.* Infant-feeding patterns and cardiovascular risk factors in young adulthood: data from five cohorts in low- and middle-income countries. *International Journal of Epidemiology* **40**, 47–62 (2011).
51. Hartwig, F. P., Loret de Mola, C., Davies, N. M., Victora, C. G. & Relton, C. L. Breastfeeding effects on DNA methylation in the offspring: A systematic literature review. *PLOS ONE* **12**, e0173070 (2017).
52. Obermann-Borst, S. A. *et al.* Duration of breastfeeding and gender are associated with methylation of the LEPTIN gene in very young children. *Pediatric Research* **74**, 344–349 (2013).
53. Naumova, O. Yu. *et al.* A Study of the Association between Breastfeeding and DNA Methylation in Peripheral Blood Cells of Infants. *Russian Journal of Genetics* **55**, 749–755 (2019).
54. Rossnerova, A. *et al.* Factors affecting the 27K DNA methylation pattern in asthmatic and healthy children from locations with various environments. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **741–742**, 18–26 (2013).
55. Sherwood, W. B. *et al.* Duration of breastfeeding is associated with leptin (LEP) DNA methylation profiles and BMI in 10-year-old children. *Clinical Epigenetics* **11**, 128 (2019).
56. Joubert, B. R. *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *The American Journal of Human Genetics* **98**, 680–696 (2016).
57. Küpers, L. K. *et al.* Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. *Nature Communications* **10**, 1893 (2019).
58. Soto-Ramírez, N. *et al.* The interaction of genetic variants and DNA methylation of the interleukin-4 receptor gene increase the risk of asthma at age 18 years. *Clinical Epigenetics* **5**, 1 (2013).
59. Tao, M.-H. *et al.* Exposures in early life: associations with DNA promoter methylation in breast tumors. *Journal of Developmental Origins of Health and Disease* **4**, 182–190 (2013).
60. Hartwig, F. P. *et al.* Association between Breastfeeding and DNA Methylation over the Life Course: Findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Nutrients* **12**, (2020).
61. Smith, A. K. *et al.* DNA extracted from saliva for methylation studies of psychiatric traits: Evidence tissue specificity and relatedness to brain. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **168**, 36–44 (2015).
62. Papavassiliou, P. *et al.* The phenotype of persons having mosaicism for trisomy 21/Down syndrome reflects the percentage of trisomic cells present in different tissues. *American Journal of Medical Genetics Part A* **149A**, 573–583 (2009).
63. Van Dongen, J. *et al.* Genome-wide analysis of DNA methylation in buccal cells: a study of monozygotic twins and mQTLs. *Epigenetics & Chromatin* **11**, 54 (2018).

64. Van Beijsterveldt, C. E. M. *et al.* The Young Netherlands Twin Register (YNTR): Longitudinal Twin and Family Studies in Over 70,000 Children. *Twin Research and Human Genetics* **16**, 252–267 (2013).

65. Boomsma, D. I. *et al.* Aggression in children: unravelling the interplay of genes and environment through (epi)genetics and metabolomics. *Journal of Pediatric and Neonatal Individualized Medicine* **4**, (2015).

66. Bartels, M. *et al.* Childhood aggression and the co-occurrence of behavioural and emotional problems: results across ages 3–16 years from multiple raters in six cohorts in the EU-ACTION project. *European Child & Adolescent Psychiatry* **27**, 1105–1121 (2018).

67. Hagenbeek, F. A. *et al.* Urinary Amine and Organic Acid Metabolites Evaluated as Markers for Childhood Aggression: The ACTION Biomarker Study. *Frontiers in Psychiatry* **11**, (2020).

68. Van Dongen, J. et al. DNA methylation signatures of aggression and closely related constructs: A meta-analysis of epigenome-wide studies across the lifespan. *Molecular psychiatry* (2021) doi:10.1038/s41380-020-00987-x.

69. Boyd, A. *et al.* Cohort Profile: The 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology* **42**, 111–127 (2013).

70. Fraser, A. *et al.* Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology* **42**, 97–110 (2013).

71. Relton, C. L. et al. Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *International journal of epidemiology* **44**, 1181–90 (2015).

72. CBS. Standard Classification of Occupations. (Heerlen/Voorburg: Central Bureau of Statistics, 2001).

73. Erikson, R., Goldthorpe, J. H. & Portocarero, L. Intergenerational class mobility and the convergence thesis: England, France and Sweden1. *The British Journal of Sociology* **61**, 185–219 (2010).

74. Dolan, C. V. *et al.* Testing Causal Effects of Maternal Smoking During Pregnancy on Offspring's Externalizing and Internalizing Behavior. *Behavior Genetics* **46**, 378–388 (2016).

75. Van Beijsterveldt, C. E. M. T. *et al.* Mode of Conception of Twin Pregnancies: Willingness to Reply to Survey Items and Comparison of Survey Data to Hospital Records. *Twin Research and Human Genetics* **11**, 349–351 (2008).

76. Odintsova, V. V *et al.* Pre- and Perinatal Characteristics Associated with Apgar Scores in a Review and in a New Study of Dutch Twins. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **22**, 164–176 (2019).

77. Meulenbelt, I., Droog, S., Trommelen, G. J., Boomsma, D. I. & Slagboom, P. E. High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *American journal of human genetics* **57**, 1252–4 (1995).

78. Sinke, L., van Iterson, M., Cats, D., Slieker, R. & Heijmans, B. DNAmArray: Streamlined workflow for the quality control, normalization, and analysis of Illumina methylation array data. (2019) doi:10.5281/ZENODO.3355292.

79. Iterson, M. Van *et al.* MethylAid : visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics* **30**, 3435–3437 (2014).

80. Van Iterson, M., Cats, D., Hop, P., BIOS Consortium & Heijmans, B. T. omicsPrint: detection of data linkage errors in multiple omics studies. *Bioinformatics* **1**, (2018).

81. Min, J. L., Hemani, G., Davey Smith, G., Relton, C. & Suderman, M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics* (Oxford, England) (2018) doi:10.1093/bioinformatics/bty476.

82. Pidsley, R. *et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology* **17**, (2016).

83. Moran, S., Arribas, C. & Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389–399 (2016).
84. Zheng, S. C. *et al.* A novel cell-type deconvolution algorithm reveals substantial contamination by immune cells in saliva, buccal and cervix. *Epigenomics* (2018) doi:10.2217/epi-2018-0037.
85. Houseman, E. A. *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86 (2012).
86. Van Iterson, M., van Zwet, E. W. & Heijmans, B. T. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biology* **18**, 19 (2017).
87. Tukey, J. W. (John W. Exploratory data analysis. *Biometric Journal* **23(4)**, 413–414 (1981).
88. Leek, J. T. & Storey, J. D. Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLoS Genetics* **3**, e161 (2007).
89. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47–e47 (2015).
90. Li, M. *et al.* EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Research* **47**, D983–D988 (2019).
91. Hartwig, F. P. *et al.* Association between breastfeeding and DNA methylation over the life course: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *bioRxiv* 800722 (2019) doi:10.1101/800722.
92. Yousefi, P. *et al.* Sex differences in DNA methylation assessed by 450 K BeadChip in newborns. *BMC Genomics* **16**, 911 (2015).
93. Husquin, L. T. *et al.* Exploring the genetic basis of human population differences in DNA methylation and their causal impact on immune gene regulation. *Genome Biology* **19**, 222 (2018).
94. Marioni, R. E. *et al.* GWAS on family history of Alzheimer's disease. *Translational Psychiatry* **8**, 99 (2018).
95. Xue, J. *et al.* Impact of vitamin D depletion during development on mouse sperm DNA methylation. *Epigenetics* **13**, 959–974 (2018).
96. Sujit, K. M. *et al.* Genome-wide differential methylation analyses identifies methylation signatures of male infertility. *Human Reproduction* **33**, 2256–2267 (2018).
97. Kim, S. K. Identification of 613 new loci associated with heel bone mineral density and a polygenic risk score for bone mineral density, osteoporosis and fracture. *PloS one* **13**, e0200785 (2018).
98. Morris, J. A. *et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nature genetics* **51**, 258–266 (2019).
99. Kemp, J. P. *et al.* Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nature genetics* **49**, 1468–1475 (2017).
100. Estrada, K. *et al.* Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nature genetics* **44**, 491–501 (2012).
101. Medina-Gomez, C. *et al.* Life-Course Genome-wide Association Study Meta-analysis of Total Body BMD and Assessment of Age-Specific Effects. *American journal of human genetics* **102**, 88–102 (2018).
102. Emilsson, V. *et al.* Co-regulatory networks of human serum proteins link genetics to disease. *Science (New York, N.Y.)* **361**, 769–773 (2018).
103. Sun, B. B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).
104. De Vries, P. S. *et al.* Multiancestry Genome-Wide Association Study of Lipid Levels Incorporating Gene-Alcohol Interactions. *American journal of epidemiology* **188**, 1033–1054 (2019).

105. Imaizumi, A. *et al.* Genetic basis for plasma amino acid concentrations based on absolute quantification: a genome-wide association study in the Japanese population. *European journal of human genetics: EJHG* **27**, 621–630 (2019).

106. Shin, S.-Y. *et al.* An atlas of genetic influences on human blood metabolites. *Nature genetics* **46**, 543–550 (2014).

107. Johnston, K. J. A. *et al.* Genome-wide association study of multisite chronic pain in UK Biobank. *PLoS genetics* **15**, e1008164 (2019).

108. Li, Y. *et al.* Genome-Wide Association Studies of Metabolites in Patients with CKD Identify Multiple Loci and Illuminate Tubular Transport Mechanisms. *Journal of the American Society of Nephrology: JASN* **29**, 1513–1524 (2018).

109. Karlsson Linnér, R. *et al.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nature genetics* **51**, 245–257 (2019).

110. Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature genetics* **51**, 237–244 (2019).

111. Hosokawa, N. *et al.* Atg101, a novel mammalian autophagy protein interacting with Atg13. *Autophagy* **5**, 973–979 (2009).

112. Ahola-Olli, A. v *et al.* Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors. *American journal of human genetics* **100**, 40–50 (2017).

113. Astle, W. J. *et al.* The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* **167**, 1415-1429.e19 (2016).

114. Rojas, D. *et al.* Prenatal Arsenic Exposure and the Epigenome: Identifying Sites of 5-methylcytosine Alterations that Predict Functional Changes in Gene Expression in Newborn Cord Blood and Subsequent Birth Outcomes. *Toxicological Sciences* **143**, 97–106 (2015).

115. McNally, K. E. *et al.* Retriever is a multiprotein complex for retromer-independent endosomal cargo recycling. *Nature Cell Biology* **19**, 1214–1225 (2017).

116. Johansson, Å., Enroth, S. & Gyllensten, U. Continuous Aging of the Human DNA Methylome Throughout the Human Lifespan. *PLoS ONE* **8**, e67378 (2013).

117. Gervin, K. *et al.* Intra-individual changes in DNA methylation not mediated by cell-type composition are correlated with aging during childhood. Clinical Epigenetics 8, 110 (2016).

118. Dunn, E. C. *et al.* Sensitive periods for the effect of childhood adversity on DNA methylation: Results from a prospective, longitudinal study. 1 doi:10.1101/271122.

119. Simpkin, A. J. *et al.* Longitudinal analysis of DNA methylation associated with birth weight and gestational age. *Human Molecular Genetics* **24**, 3752–3763 (2015).

120. Agha, G. *et al.* Birth weight-for-gestational age is associated with DNA methylation at birth and in childhood. *Clinical Epigenetics* **8**, 118 (2016).

121. Alfano, R. *et al.* 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**, 30–44 (2019).

122. Richmond, R. C. *et al.* Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Human Molecular Genetics* **24**, 2201–2217 (2015).

123. Tobi, E. W. *et al.* DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Human Molecular Genetics* **18**, 4046–4053 (2009).

124. Muniz, L. C., Menezes, A. M. B., Buffarini, R., Wehrmeister, F. C. & Assunção, M. C. F. Effect of breastfeeding on bone mass from childhood to adulthood: a systematic review of the literature. *International breastfeeding journal* **10**, 31 (2015).

125. Dessì, A. *et al.* Exploring the Role of Different Neonatal Nutrition Regimens during the First Week of Life by Urinary GC-MS Metabolomics. *International Journal of Molecular Sciences* **17**, 265 (2016).

126. San-Cristobal, R. *et al.* Gene methylation parallelisms between peripheral blood cells and oral mucosa samples in relation to overweight. *Journal of Physiology and Biochemistry* **73**, 465–474 (2016).

127. Van Dongen, J. *et al.* Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nature Communications* **7**, 11115 (2016).

128. Colodro-Conde, L., Sánchez-Romera, J. F. & Ordoñana, J. R. Heritability of Initiation and Duration of Breastfeeding Behavior. *Twin Research and Human Genetics* **16**, 575–580 (2013).

129. Merjonen, P., Dolan, C. v., Bartels, M. & Boomsma, D. I. Does Breastfeeding Behavior Run in Families? Evidence From Twins, Their Sisters and Their Mothers in the Netherlands. *Twin Research and Human Genetics* **18**, 179–187 (2015).

130. Eriksson, N. *et al.* Genetic variants associated with breast size also influence breast cancer risk. *BMC Medical Genetics* **13**, 53 (2012).

131. Cunningham, B., Bass, S., Fuh, G. & Wells, J. Zinc mediation of the binding of human growth hormone to the human prolactin receptor. *Science* **250**, 1709–1712 (1990).

132. Sukonina, V. *et al.* FOXK1 and FOXK2 regulate aerobic glycolysis. *Nature* **566**, 279–283 (2019).

133. Hannenhalli, S. & Kaestner, K. H. The evolution of Fox genes and their role in development and disease. Nature reviews. *Genetics* **10**, 233–40 (2009).

134. Tachmazidou, I. *et al.* Whole-Genome Sequencing Coupled to Imputation Discovers Genetic Signals for Anthropometric Traits. *American journal of human genetics* **100**, 865–884 (2017).

135. Crujeiras, A. B. *et al.* An Epigenetic Signature in Adipose Tissue Is Linked to Nicotinamide N-Methyltransferase Gene Expression. *Molecular Nutrition & Food Research* **62**, 1700933 (2018).

136. Bardanzellu, F., Fanos, V., Strigini, F. A. L., Artini, P. G. & Peroni, D. G. Human Breast Milk: Exploring the Linking Ring Among Emerging Components. *Frontiers in Pediatrics* **6**, 215 (2018).

137. Lehmann, G. M. *et al.* Environmental Chemicals in Breast Milk and Formula: Exposure and Risk Assessment Implications. *Environmental Health Perspectives* **126**, 096001 (2018).

138. Gómez-Gallego, C. *et al.* Human Breast Milk NMR Metabolomic Profile across Specific Geographical Locations and Its Association with the Milk Microbiota. *Nutrients* **10**, 1355 (2018).

139. Hermansson, H. *et al.* Breast Milk Microbiota Is Shaped by Mode of Delivery and Intrapartum Antibiotic Exposure. *Frontiers in Nutrition* **6**, 4 (2019).

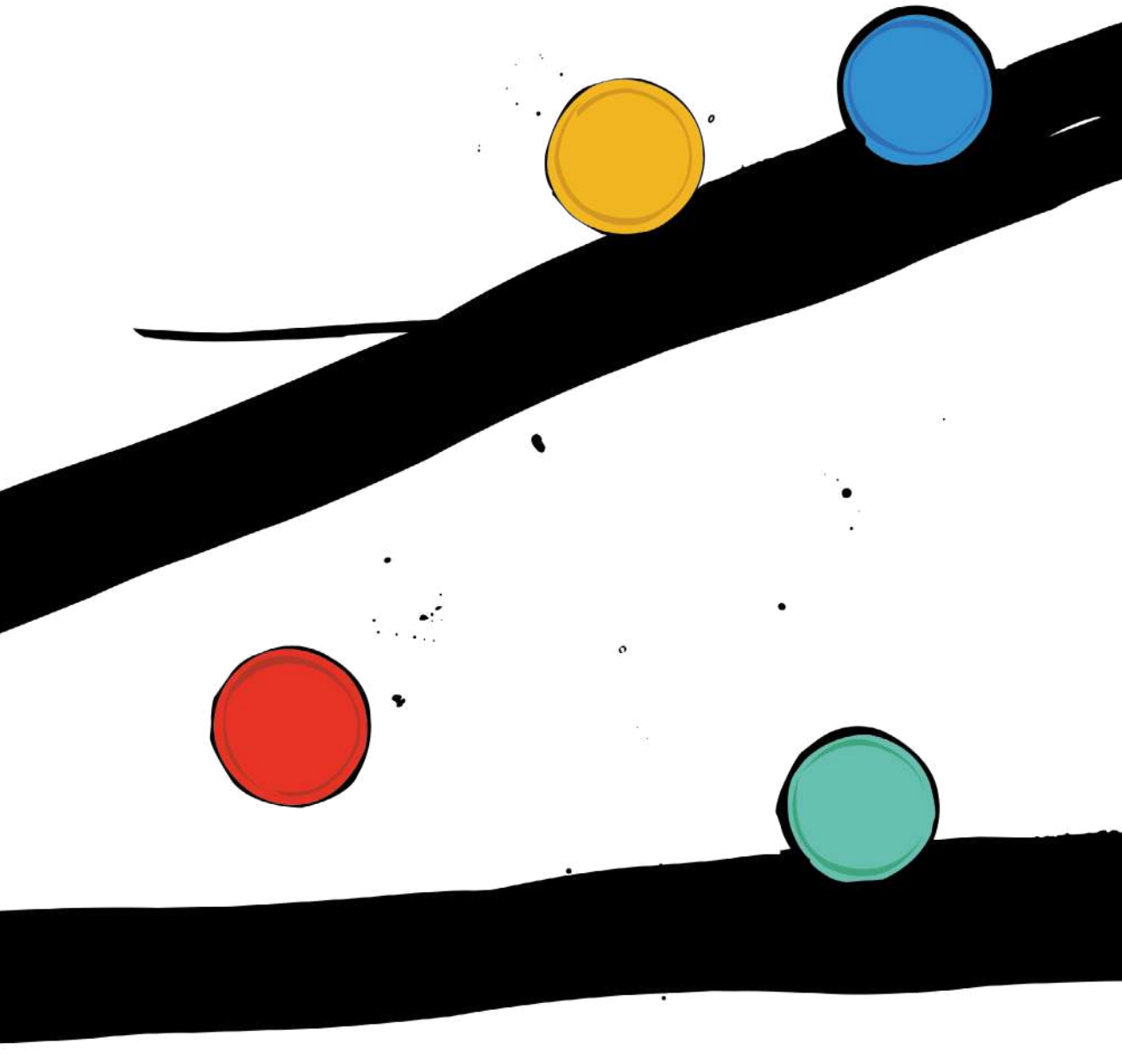
Supplements

The supplementary materials are available online at

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6893543/

Chapter 5

Handedness and 23 early life characteristics: a review and a study in 37,495 Dutch twins



Abstract

In singletons a range of early life characteristics is associated with handedness, but several associations remain disputed, which in part may be due to the inclusion or exclusion of mixed-handed individuals. We carried out a literature review of large association studies of handedness with early life characteristics and examined associations between 23 early life characteristics and handedness at 5 years in a large sample of twins. Three definitions of handedness were considered: right- versus left-handed (*N*=36,997), right- versus mixed-handed (*N*=31,923), and right-handed versus all others (*N*=37,495). Our main aim was to test if the associations with sex, birth weight, gestational age, and season of birth, as established in singletons replicate in twins, and to test twin-specific variables, such as zygosity, chorionicity, birth order, and intertwin delivery time. Sex and gestational age were associated with handedness regardless of the definition. Left-handedness and non-right-handedness were associated with mother's and father's left-handedness. Left-handedness was associated with breastfeeding. Mixed-handedness was associated with neurodevelopmental delay and higher externalizing problems later in childhood. Other previously reported associations with handedness were not replicated, and no twin-specific characteristic was related to handedness.

Published as: Odintsova V., Dolan CV, Beijstervaldt T., Ligthart L., Willemsen G., de Geus E.J.C., van Dongen J., Boomsma DI. (2022) Handedness and 23 early life characteristics: a review and a study in 37,495 Dutch twins. *Laterality* (in revision).

Introduction

Handedness refers to the preference or use of one hand over the other hand and may be observed on ultrasound scans already early in prenatal life[1–3]. The prevalence of right-handedness is around 90% in nearly all human populations, with the remaining proportion of individuals being left- or mixed-handed, which is also referred to non-right-handedness.

Etiological hypotheses of handedness consider genetic and epigenetic explanations, the influence of environmental factors, and stochastic events Medland et al.[4] analyzed handedness data from 54,270 twins and their non-twin siblings in 25,732 Australian and Dutch families. The study revealed no differences between twins and their non-twin siblings, no evidence of prenatal hormone transfer in twin pairs of opposite sex, of mirror imaging, of twin specific effects, or of differences in prevalence between zygosity groups. After correction for year of birth, birthweight, and sex, heritability accounted for around 24% of the variance in handedness, with the remainder accounted for by non-shared environmental influences, which may include stochastic events[5,6]. According to Annett's right-shift theory[5] and McManus' dextra-chance theory[6], a person becomes either left-handed, mixed-handed, or right-handed by chance due to a random shift away from strong right-handedness towards weak right-handedness (mixed-handedness) or left-handedness. Identification of the environmental factors is likely to be difficult, and McManus[7] predicted that "few environmental factors accounting for more than a percent or two of variance in handedness" will be found. A recent epigenetic study found that epigenetic mechanisms are involved in handedness and reported two DNA regions with differential DNA methylation associated with left-handedness, but the variance explained by DNA methylation also is low[8] (see **Chapter 6**).

In research into etiological factors, many associations with pre- and perinatal factors have been included with multiple studies linking early life factors to handedness. For example, Bakan et al.[9] related non-right-handedness to "birth stress" as indexed by multiple birth, premature birth, prolonged labor, caesarian section, breach position, and breathing difficulty at birth, which all may cause perinatal hypoxia and brain damage. One hypothesis related the origins of handedness to prenatal levels of testosterone and suggested that left-handed individuals may have been exposed to higher levels of testosterone in comparison with right-handed, thereby explaining the higher prevalence of left-handedness in males[10,11]. However, later studies demonstrated the opposite that prenatal level of testosterone could be related to right-handedness[12–14].

We carried out a review of large association studies of handedness with early life characteristics (**Table 1**). The largest study of early life characteristics associated with handedness was performed in the UK Biobank cohort of ~500,000 adults[15]. The study defined handedness as being left-handed versus right-handed as definition of handedness, and reported associations with sex (more left-handed males), year and location of birth (left-handers increased with year of birth up to 1970, and more in England than outside England and UK), season of birth (more left-handed persons born in summer months), birthweight (more left-handedness given lower birthweight), being part of a multiple birth (more left-handed individuals from twin pairs), and breastfeeding (more left-handed persons not being breastfed).

An early meta-analysis of 23 large-scale studies including total 46,699 individuals found that that prenatal stress, fetal presentation (breech position in males), gestational age, and mode of delivery (caesarean section) were associated with non-righthandedness[16]. A Danish study of 35,206 singletons found associations with mixed-handedness for mother's and father's handedness, gestational age (being preterm), and mode of conception (children conceived after intrauterine insemination were twice the risk of being mixed-handed compared to naturally conceived children)[17]. A study of 1,031 school students at around 16 years old showed an association of low Apgar scores with non-right-handedness[18]. For parental age, findings were inconsistent with some authors reporting an association with offspring left-handedness[19,20], and others reporting no association[16,18,21]. Stressful events in the third period of pregnancy were related to a higher prevalence of mixed-handedness in offspring[22,23].

Regarding gestational age, a meta-analysis of 18 studies, which included total 10,117 individuals, reported 22% non-righthanded in preterm born compared to 12% in controls[24]. The larger proportion of non-righthanded (23%) was also reported in extreme preterm births[25].

Handedness has been related to neurodevelopment and neuropathology[26], and externalizing problems later in life. Higher externalizing problems, as indexed by aggression were associated with left-handedness (measured by an online tapping task) in study of 305 undergraduate students[27] and in an internet survey of 20,539 adults[28]. A study of psychiatric disorders in 692 children yielded 78% increased odds of oppositional defiant disorder for left-handed children[29]. Non-right-handed adolescents had higher rates of behavioral problems in Dutch study of 2,096 adolescents. However, externalizing problems showed no association with non-right-handedness in this study[30].

A higher incidence of left-handedness in twins and triplets, compared to singletons, has been reported in some studies[15,31–36], but not in all[4].

Monozygotic (MZ) and dizygotic (DZ) twins did not differ in left-handedness prevalence[35,37,38], nor were differences seen as a function of chorionicity[37,39]. Some studies reported evidence that first-born twins are more likely to be left-handed than second-born twins[40], but this was not replicated in other studies[14,37]. Lower prevalence of left-handedness was observed in female twins in opposite-sex pairs, compared to female twins in same-sex pairs[14]. Twins are frequently born preterm and with low birthweight and are at greater risk of certain congenital disorders and perinatal morbidity[41].

Our review of previous studies shows inconsistencies in results, with also pointing to a possible explanation as studies have applied different definitions of, and different methods to assess, handedness. Some, but not all, studies discard mixed-handed individuals, and meta-analyses have integrated studies with different assessment methods. The present study, therefore, include three definitions: right-handed (RH) vs left-handed (LH), right-handed versus mixed-handed (MH), and right-handed vs non-right-handed (NRH). With respect to these three definitions, we aim to: (1) to determine whether the risk factors for handedness, as previously identified in singleton and population-based cohorts, replicate in twins, and (2) to determine the associations between handedness and twin-specific characteristics. To these ends, we analyzed data on handedness from 5-year-old twins

registered in the Netherlands Twin Register (NTR) who were followed longitudinally since birth[42]. The twins were born between 1989 and 2014. All analyses were performed with generalized equation estimates to account for relatedness between twins.

Methods

Overview. We studied the association between handedness at 5 years old and early life characteristics and outcomes in twins. First, we perfomed a literature review to compile a list of possibly relevant early life characteristics for analysis. Second, we perfomed an association study in twins. The NTR includes 38,496 5-year-old twins, who were born between 1989 and 2014, and were registered as newborns with the NTR by their parents. Of these, 97% had information on drawing hand preference at 5 years old, and 3% had no information on handedness. After exclusion of the twins (3%) with no data on handedness, sample size was 37,495 twins in 18,630 complete and 235 incomplete twin pairs. As mentioned above, three handedness definitions were considered: left-handedness (LH, left-handed vs. RH, right-handed), mixed-handedness (MH, both hands vs. right-handed), and non-right-handedness (NRH, left- and mixed-handed vs. right-handed). Twenty three variables related to prenatal and early life were studied in association with handedness (see **Supplementary Table 1**): six that were also tested in the currently largest study, performed in UK Biobank, by de Kovel et al. (2019)[15], and 13 characteristics available in the NTR, which featured in previous studies of early life characteristics in association with handedness, including such twin-specific characteristics as zygosity and chorionicity. In addition, we included amnionicity and time interval between the birth of the 1st and 2nd twin that have not been considered previously. As handedness is related to neurodevelopmental delay and externalizing problems, we included neurodevelopmental delay measured by bowel and bladder toilet skill delay assessed at age 5. Children with typical development usually acquire these skills, assessed having less than four wetting accidents per week, by the age of four[43,44]. Externalizing problems, specifically aggression, was measured in this cohort by the the CBCL (Achenbach) when twins reached age 7.

Literature review. To identify potentially early life characteristics to include, we reviewed the literature. We conducted a literature search in PubMed based key words with further semi-automized selection based on abstract screening with ASReview tool (https://asreview.nl, developers: Utrecht University). Specifically, we carried out the searches of literature in PubMed with key words “handedness” (62,901 results, of them 10,000 uploaded; date 9.05.2021) and “handedness&twins” (290 results; date 9.05.2021). The further selection was performed based on abstract screening with ASReview tool of the combined list of references with duplicates removed to select the papers with studies on association studies of handedness with early life characteristics (prenatal, perinatal, and early life characteristics). The study selection criteria included the criterion of sample sizes greater than 200 individuals. However, we included two clinical studies with N<200 because they addressed prenatal stress and gestational age in preterm population, i.e., characteristics which have hardly been considered in the literature. If a recent meta-analysis was available, we discarded other population-based studies if they concerned

TABLE 1.
Overview of the association studies of handedness and early life characteristics

First author, year	Sample size	Age of handedness measurement	Handedness phenotype	Handedness measurement	Characteris-tics asso-ciated with handedness	Characte-ristics not associated with handedness
Mixed twin-singleton population studies						
De Kovel, 2019	421,776	40-69 years	RH/LH	Self-report	Being multiple, year of birth, sex, season of birth, birth-weight, being breastfed	Maternal smoking
Vuoksimaa, 2009	30,161	18-69 years	LH in RH/LH/MH	Self-report on childhood and current hand preference	Being multiple, sex	Maternal smoking
			MH in RH/LH/MH		Sex; higher MH in triplets vs singletons and twins	Zygosity, birth order
					Birthorder	Zygosity, birth order
Dra-govic, 2013	1,031	Mean 16.2	RH/NRH	Self-report	Maternal smoking, low Apgar scores	Being multiple, birth order, parental age, mode of delivery
Singleton studies						
Hujoel, 2019	62,129	Children	RH/NRH	Different measure-ments in meta-analysis	Being breastfed	None
Searleman, 1989	46,699	Children-adults (meta-analysis)	RH/NRH	Different measure-ments in meta-analysis	Mode of delivery, gestational age, sex, birthweight, fetal presen-tation, maternal smoking, birth order position in family	Mother's age at birth, being multiple
Zhu, 2009	35,206	7 years	RH/NRH	Report on "Which hand does your child use most?" or "Is your child right-handed or left-handed?" (3 response categories: RH, LH, both) and Annett Hand Preference Questionnaire	Gestational age, parental handedness, mode of conception, maternal smoking, contraception during 1st trimester	None

Table 1 (continued)

First author, year	Sample size	Age of handedness measurement	Handedness phenotype	Handedness measurement	Characteristics associated with handedness	Characteristics not associated with handedness
Denny, 2012	21,847	7 years	RH/NRH (LH+MH)	Mother's survey	Being breastfed	
Van der Feen, 2020	20,539	Mean 41.3	Continuous	Hand skill assessed by alternating key press task ("tapping task")	Aggression score	None
John-ston, 2013	12,686	Mean 7.4	RH/LH	Question on writing hand	Being breastfed, sex, parental age at birth, birthweight	Mode of delivery, gestational age, maternal smoking, socio-economic status
Domel-lof, 2011	10,117	3-19 years	RH/NRH	Different measurements in meta-analysis	Gestational age, neurological and neuro-psychological outcomes	None
BaileyMc Keever, 2004	2,151	Undergraduate students	RH/LH	Writing hand	Mother's age at birth	List of 25 factors, including mode of delivery, gestational age, birth-weight, fetal presentation
Van der Hoorn, 2010	2,096	Mean 13.6	RH/NRH	Report on "What is the hand you are writing with?" (3 response categories: RH, LH, alternating)	Thought problems, social problems, being withdrawn and depressed (psychotic items)	Obstetric factors (gestational age, birth-weight, caesarean section, vacuum or forceps assisted birth), treated with oxygen or incubator as neonate, externalizing problems
Sutcliffe, 2005	1,525	5 years	Continuous	McCarthy Scale of children's abilities (motor scale), parental report: child handedness for drawing and writing	None	Parental handedness, mode of conception

Table 1 (continued)

First author, year	Sample size	Age of handedness measurement	Handedness phenotype	Handedness measurement	Characteristics associated with handedness	Characteristics not associated with handedness
Fagard, 2021	1,129	5 years	Continuous and RH/LH	Handedness index based on 8-item hand preference test (from LH to RH) and absolute handedness index (from non-lateralized to non-late-ralized). Binomial variable based on handedness index: RH vs LH	Father's handedness, season of birth, gestational age, fetal presentation (breech), being breastfed	Sex, mother's handedness
Karev, 2008	870	Mean 16.6	RH/MH/LH	Drawing hand (Chapman and Chapman's inventory)	None	Parental age at birth, season of birth
Obel, 2003	824	3 years	RH/MH	Mixed-handedness based on maternal report at 3 years old (use of hand in 5 activities)	Prenatal stress in 3d trimester	None
Logue, 2015	692	4-18 years (clinical setting, predominantly African-American ancestry)	RH/MH	Writing hand	Psychiatric diagnosis including oppositional defiant disorder	None
Dinsdale, 2011	395	Mean 19.2	Continuous	Edinburgh Inventory of Handedness	Aggression score	Sex
van der Elst, 2021	294	5.67-15.08 years	RH/LH (LH+MH)	Lateral preference (hand, foot, eye, ear preference)	None	Sex, fetal presentation, mode of delivery
Marlow, 2007	241	6 years	Continuous	Neuropsycho-logical battery (fingertip tapping)	Gestational age	None
Van Heerwaarde, 2020	179	5 years	RH/NRH (MH and LH combined)	Movement Assessment Battery for Children second edition, Dutch version (MABC-2-NL): "the hand used to write or draw with" at school age in preterm clinical group (<28 weeks GA)	Parental handedness, gestational age	Being multiple, sex, season of birth, birthweight, Apgar scores, parental education level

Table 1 (continued)

First author, year	Sample size	Age of handedness measurement	Handedness phenotype	Handedness measurement	Characteris-tics asso-ciated with handedness	Characte-ristics not associated with handedness
Gutteling, 2007	110	6 years	RH/MH	Hand preference based on 8 activities assessed by independent observers	Gestational age, mother's handedness, prenatal stress	None
Sicotte, 1999	19,938	Children-adults (meta-analysis)	RH/LH	Different measure-ments in meta-analysis	Being multiple	Zygosity
van Beijster-veldt, 2016	1,8222	5 years	RH/LH/MH	Parental report about which hand is used for drawing at the survey of age 5	None	Chorionicity
Medland, 2003	14,838	Children-adults (meta-analysis)	RH/LH	Self-report on "Which hand would you use to write a letter?" and "Which hand would you use to throw a ball to hit a target"	None	Birth order, being multiple, zygosity
Vuoksimaa, 2010	4,736	14 years	RH/LH/MH	Self-report	Sex	Birth order
Ooki, 2006	4,164	1-15 years	RH/LH	Parental report on "Which hand would your twin children predominantly use, if pos-sible, to write a letter?"	None	Being multiple
Orlebeke, 1996	3,400	Mean 17.8	RH/LH	Report on "Do you consider yourself predominantly right-handed or predominantly left-handed?"	Birth order, being multiple, zygosity	None
Heikkila, 2015	2,252	Mean 12 [1-36 years]	RH/LH	Parental report and self-report	Birthweight (in triplets)	None
Derom, 1996	1,616	6 - 28 years	RH/LH	Parental report	Being multiple	Sex, birth order, chorionicity, zygosity

Table 1 (continued)

First author, year	Sample size	Age of handedness measurement	Handedness phenotype	Handedness measurement	Characteris-tics asso-ciated with handedness	Characte-ristics not associated with handedness
Elkadi, 1999	1,476	Mean 23.5	Continuous	Survey on hand use for 3 activities	None	Birth order
James& Orlebeke, 2002	606	Mean 17.8	RH/LH	Report on "Do you consider yourself predominantly right-handed or predominantly left-handed?"	Birth order	None

Abbreviations: LH, left-handedness; RH, right-handedness, MH, mixed-handedness, NRH, non-righthandedness, Continuous, score measurement. The studies in the table are ordered by sample size in three groups: mixed twin-singleton population studies, singleton population studies, and twin studies.

the same characteristics and used the same definition of handedness. To determine the association of handedness with aggression, we conducted an additional literature search using the term "handedness&aggression" (272 results; date 11.05.2021), repeated the selection, and included three studies on aggression and handedness. One study with findings on oppositional defiant disorder was added from reference list. Thus, 30 studies were included in the final analysis. All studies were evaluated by the following criteria: year of publications, sample size, type of study (population-based (mixed singleton-twin)/ singleton/twin study), handedness definition, and association of handedness with early life characteristics.

Subjects. The NTR recruits families with twins weeks to months after birth. Informed consent was obtained from parents. Surveys, including questions on pregnancy, birth and outcomes, were sent to mothers after registration of newborn twins. The zygosity of same-sex twins was determined by blood/DNA polymorphisms (18%), or by parental report based on a questionnaire, addressing several aspects of physical resemblance and the degree to which the twins are confused by parents, other relatives and strangers (see Ligthart et al. 2019 for a detailed description). In a small number of twins the information on zygosity was missing (32 cases, 0.1%). There were 55 families with 2 twin pairs, of which three families included two MZ twin pairs, 37 families included two DZ twin pairs, and 15 families included a MZ and a DZ twin pair.

Handedness. Handedness was assessed at age 5 years based on the answer on question: "Which hand do the children use to draw on paper?" reported for 1st and 2nd born twin. The answers were "left hand", "right hand", "both hands", and "I don't know". The cases with replies "I don't know" (*N*=52) and missed data (*N*=949) were excluded from the study, and the final variable was coded left-handed, right-handed, and both hands.

Early life characteristics. Based on the literature review, we tested for the effects of five sets of early life characteristics (see **Supplementary Table 1**): 1) *General characteristics:* sex (males, females), year of birth, mother's and father's handedness; 2) *Prenatal characteristics:* mother's and father's age at birth, mode of conception (Spontaneous/ Artificial), prenatal maternal smoking (no/yes), maternal stress during pregnancy (no/yes); 3) *Perinatal characteristics:* season of birth (being born in summer or not), gestational age (continous), fetal presentation at birth (cephalic presentation and non-cephalic presentation: breech and horizontal), mode of delivery (vaginal and intervention with vacuum extraction, forceps or cesarean section), birthweight (continous), Apgar score at 1st minute (continuous); 4) *Early life postnatal characteristics:* breastfeeding (no/yes), neurodevelopmental delay (no/yes) assessed based on bowel and bladder toilet skills delay at age 5 (reports on questions "How often do the children defecate in their pants?" and "How often do the children pee in their pants during the day?"; see **Appendix 1**), and externalizing problems, specifically aggression, using the subscale Aggressive Behavior of the Child Behavior Checklist (CBCL), which the mother completed when children were 7 years old[45]; and 5) *Twin-specific characteristics:* zygosity (dizygotic/monozygotic), chorionicity (dichorionic/monochorionic), amnionicity (diamniotic/monoamniotic), birth order (first/second born), and time interval between the birth of the 1st and 2nd twin. Thus, 23 early life characteristics were tested for association with left-handedness (15

categorical and 8 continuous). Pre- and perinatal characteristics were obtained from a survey that was sent to mothers within the first year after birth. Handedness of the parents and information on breastfeeding were obtained from a survey collected when the twins are 2 years of age[46] that included the question "Are the following persons predominantly left-handed or right-handed (mother, father)?" (three answer categories were left-handed, right-handed, and both hands), and the question "Did your twin get breastfeeding?" (answer categories "no", "less than 2 weeks", "2 to 6 weeks", "6 weeks to 3 months", "3 to 6 months", "more than 6 months" were categorized in yes/no). The information on toilet skills was obtained from the same survey that addressed children handedness, at 5 years old. Information on chorion status was obtained by linkage to the records from the database of the Dutch pathological anatomy national automated archive (PALGA)[39]. The time interval between the twin births included 6 cases with extreme values (i.e., more than 33 hours of time between birth of 1st and 2nd born). These extreme values were recoded as missing. There were >45% missing data due to absence of characteristics in earlier surveys, no response given, or absence of data in linkage records. The questions on several early life characteristics were included later to the surveys, and that explain missing data: aggression score at 7 years (45%), chorionicity and amnionicity (70%), fetal presentation at birth (81%), maternal stress during pregnancy (81%), and Apgar score at 1 minute (86%). The distributions of the characteristics in LH, RH and MH groups are presented in **Supplementary Figures 1 and 2**.

Statistical analysis. All statistical analyses were performed in R version 4.1.0. Due to missingness, the number of subjects in the regression analyses of individuals varied between 4,515 and 37,495 subjects.

Descriptive statistics. The frequencies of categorical variables, and means and standard deviations (SD) of continuous variables were obtained in total sample, and within the LH, RH, and MH groups. As descriptive measures of association, we report Pearson correlations for continuous variables (R package "stats"), polychoric correlations for categorical variables (R package "polychor"), and point polyserial correlations for continuous with categorical variables (R package "stats").

Association analysis: For testing associations between handedness and characteristics, we used generalized estimating equation (GEE) logistic models to account for relatedness of twins. We conducted 23x3 individual regression analyses, in which we regressed handedness on the 23 predictors. The coding of variables included in the models is presented in **Supplementary Table 1**. Given dichotomous predictors and dichotomized continuous predictors, we report odds ratio and 95% confidence intervals, which were based on the GEE logistic regression output.

Multiple testing correction. We set the family-wise alpha to equal .05 and applied the Bonferroni correction (per dependent variable: .05/23 = .002). As this correction for multiple testing can be conservative, we also considered the correction based on Nyholt (2004), which involves calculating the effective number of test, given the correlations between the predictors. Correlations between 23 early life characteristics are presented in **Supplementary Figure 3** and **Supplementary Table 2**. The Nyholt procedure produced approximately the same per-test-alpha of .002 (which is consistent with the low average

correlation of .14 amongst the predictors). Given the two definitions (left-handedness and mixed-handedness; note that non-right-handedness is the combination of these two), we set the the alpha-per-test to equal $\alpha = 0.05/(22.5*2) = 0.0011$.

Results

Literature review. The reviewed studies were published between 1989 and 2021. The sample sizes ranged from 240 to 421,776, except for two studies in less than 200 singletons in clinical settings[22,25]. Of 29 reviewed studies, there were two studies with mixed singleton-twin samples, 19 studies of singletons, and 10 studies of twins and triplets. The studies analyzed different definitions of handedness: right-handed vs left-handed[14,15,19–21,24,25,29,33,35–40,46–48], right-handed vs mixed-handed[17,21–23], or combined left-handed with mixed-handed in non-right-handed group[16,18,24,25,30,49–51], and handedness preference score as a continuous trait[28,52–54]. Description of the studies and detected associations are summarized in **Table 1**. Our selection of the early life characteristics and expected associations based on the literature are presented in **Supplementary Table 1**. We investigated left-handedness (LH), mixed-handedness (MH) and non-right-handedness (NRH) as the binary traits to compare it with previous studies performed with different handedness phenotypes.

Descriptives: prevalence of handedness in twin families. In the parental generation with known handedness (*N*= 31,813, 49.9% males) 84.7% were right-handed, 11.4% were left-handed and 3.6% mixed-handed. The study cohort included 37,495 twins from 18,850 twin pairs born between 1986 and 2014, 49.7% males and 35% MZ. Characteristics of the participants are presented in **Table 2**. The prevalence of right-handedness (RH) was 83.8%, left-handedness (LH) was 14.9%, and mixed-handedness (MH) was 1.3%. The prevalence of left-handedness in twins was higher than in parental generation (14.9% vs 11.4%, p<0.001) and in meta-analysis of adult healthy population (14.9% vs 10.6% 55). The prevalence of mixed-handedness in twins was lower than in parental generation (1.3% vs 3.6%, p<0.001). The prevalence of left-handedness and mixed-handedness was higher in males than in females (15.9% vs 13.8% for left-handedness, p<0.001; 1.8% vs. 0.8% for mixed-handedness, p<0.001). In comparison with right-handed twins, left-handed twins more often had left-handed mothers (14.7% vs 9.8%; p<0.001) and fathers (14.6% vs 11.9%; p<0.001), one or both not-right-handed parents (33.3% vs. 26.9%; p<0.001). No differences were observed between right-handed and mixed-handed twins (one or both parents were non-right-handed in 31.7% of mixed-handed twins vs 26.9% in right-handed).

Descriptives: correlations with handedness. Correlations between life characteristics were mostly low (r <.5) except several strong correlations (r > .6 between amnionicity, chorionicity and zygosity, between mode of conception and zygosity, between gestational age and birthweight, between mother's and father's age) (see **Supplementary Figure 3** and **Table 2**). Correlations between handedness and early life characteristics ranged between -.037 and .139 for left-handedness definition, between -.190 and .191 for mixed-handedness definition, and between -.035 and .119 for non-right-handedness

TABLE 2.
Characteristics of participants

		N total	Total *N*=37,495	RH *N*=31,425	LH *N*=5,572	MH *N*=498
General characteristics						
Range of year of birth			1986-2014	1986-2014	1986-2014	1987-2014
Sex		37,495	18616 (49.6%)	15314 (48.7%)	2964 (53.2%)	338 (67.9%)
	Males, n(%)					
	Females, n(%)		18879 (50.4%)	16111 (51.3%)	2608 (46.8)	160 (32.1%)
		31,883				
Mother's handedness	*LH, n(%)*		3357 (10.5%)	2622 (9.8%)	699 (14.7%)	36 (8.8%)
	RH, n(%)		27379 (85.9%)	23170 (86.7%)	3858 (81.3%)	351 (85.4%)
	MH, n(%)		1133 (3.6%)	920 (3.4%)	189 (4%)	24 (5.8%)
		31,777				
Father's handedness	*LH, n(%)*		3931 (12.4%)	3185 (12%)	695 (14.7%)	51 (12.5%)
	RH, n(%)		26536 (83.5%)	22358 (84%)	3850 (81.2%)	328 (80.2%)
	MH, n(%)		1176 (3.7%)	971 (3.6%)	177 (3.7%)	28 (6.8%)
Prenatal characteristics						
Mother's age at birth, mean(sd)		36,364	31.4 (4)	31.4 (4)	31.4 (3.9)	31.4 (4.3)
Father's age at birth, mean(sd)		35,990	33.9 (4.7)	33.9 (4.7)	33.8 (4.7)	34.1 (5.2)
Mode of conception		34,526				
	Spontaneous, n(%)		25675 (74.4%)	21457 (74.2%)	3873 (75.3%)	345 (73.7%)
	Artificial, n(%)		8851 (25.6%)	7455 (25.8%)	1273 (24.7%)	123 (26.3%)
Prenatal maternal smoking		36,360				
	No, n(%)		30000 (82.5%)	25206 (82.7%)	4416 (81.6%)	378 (78.4%)
	Yes, n(%)		6360 (17.5%)	5257 (17.3%)	999 (18.4%)	104 (21.6%)
Maternal stress during		7,202				
pregnancy	*No, n(%)*		4540 (63%)	3837 (63.4%)	651 (61%)	52 (61.2%)
	Yes, n(%)		2662 (37%)	2212 (36.6%)	417 (39%)	33 (38.8%)

Table 2 (continued)

		N total	Total *N*=37,495	RH *N*=31,425	LH *N*=5,572	MH *N*=498
Mother's age at birth, mean(sd)		36,364	31.4 (4)	31.4 (4)	31.4 (3.9)	31.4 (4.3)
Father's age at birth, mean(sd)		35,990	33.9 (4.7)	33.9 (4.7)	33.8 (4.7)	34.1 (5.2)
Mode of conception		34,526				
	Spontaneous, n(%)		25675 (74.4%)	21457 (74.2%)	3873 (75.3%)	345 (73.7%)
	Artificial, n(%)		8851 (25.6%)	7455 (25.8%)	1273 (24.7%)	123 (26.3%)
Prenatal maternal smoking		36,360				
	No, n(%)		30000 (82.5%)	25206 (82.7%)	4416 (81.6%)	378 (78.4%)
	Yes, n(%)		6360 (17.5%)	5257 (17.3%)	999 (18.4%)	104 (21.6%)
Maternal stress during pregnancy		7,202				
	No, n(%)		4540 (63%)	3837 (63.4%)	651 (61%)	52 (61.2%)
	Yes, n(%)		2662 (37%)	2212 (36.6%)	417 (39%)	33 (38.8%)
Perinatal characteristics						
Season of birth		37,491				
	Not Summer, n(%)		8896 (23.7%)	7497 (23.9%)	1270 (22.8%)	129 (25.9%)
	Summer, n(%)		9707 (25.9%)	8082 (25.7%)	1490 (26.7%)	135 (27.1%)
Fetal presentation		7,268				
	Head, n(%)		4769 (65.6%)	4003 (65.6%)	718 (66.5%)	48 (55.8%)
	Breech, n(%)		1923 (26.5%)	1617 (26.5%)	283 (26. %2)	23 (26.7%)
	Horizontal, n(%)		576 (7.9%)	483 (7.9%)	78 (7.2%)	15 (17.4%)
Mode of delivery		31,818				
	Vaginal, n(%)		18594 (58.4%)	15549 (58.4%)	2789 (59.0%)	256 (57.5%)
	Caesarean session planned, n(%)		3861 (12.1%)	3259 (12.2%)	555 (11.7%)	47 (10.6%)
	Vacuum/Forceps, n(%)		2942 (9.2%)	2432 (9.1%)	470 (9.9%)	40 (9.0%)
	Caesarean session urgent, n(%)		6421 (20.2%)	5405 (20.3%)	914 (19.3%)	102 (22.9%)
Gestational age, mean(sd)		36,277	36.5 (2.5)	36.6 (2.5)	36.4 (2.6)	36.1 (2.8)
Birth weight, mean(sd)		36,045	2490.3 (549.8)	2493.5 (546)	2480.2 (566.1)	2404.9 (596)
Apgar score at 1 minute, mean(sd)		5,281	8.3 (1.8)	8.3 (1.7)	8.3 (1.7)	7.7 (2.5)

Table 2 (continued)

		N total	Total *N*=37,495	RH *N*=31,425	LH *N*=5,572	MH *N*=498
Postnatal characteristics						
Breastfeeding		32,038				
	No, n(%)		13057 (40.8%)	10835 (40.4%)	2053 (43%)	169 (40.8%)
	Yes, n(%)		18981 (59.2%)	16012 (59.6%)	2724 (57%)	245 (59.2%)
Neurodevelopmental delay at 5 years		37,106				
	No, n(%)		36252 (97.7%)	30408 (97.8%)	5381 (97.7%)	463 (94.3%)
	Yes, n(%)		854 (2.3%)	700 (2.25%)	126 (2.3%)	28 (5.7%)
Aggression score at 7 years, mean(sd)		20,595	5.1 (4.9)	5.1 (4.9)	5.2 (4.8)	7.1 (5.6)
Twin-specific characteristics						
Birthorder		37,465				
	1st born, n(%)		18776 (50.1%)	15684 (49.9%)	2863 (51.4%)	229 (46%)
	2nd born, n(%)		18689 (49.8%)	15715 (50%)	2706 (48.6%)	268 (53.8%)
Zygosity		37,321				
	Monozygotic, n(%)		13129 (35.2%)	10956 (35%)	2013 (36.3%)	160 (32.4%)
	Dizygotic, n(%)		24192 (64.8%)	20327 (65%)	3531 (63.7%)	334 (67.6%)
Chorionicity		11,201				
	Monochorionic, n(%)		3483 (31.1%)	2908 (31%)	533 (32.2%)	42 (26.6%)
	Dichorionic, n(%)		7718 (68.9%)	6478 (69%)	1124 (67.8%)	116 (73.4%)
Amnionicity		11,155				
	Monoamniotic, n(%)		281 (2.5%)	231 (2.5%)	49 (3%)	1 (0.6%)
	Diamniotic, n(%)		10874 (97.5%)	9120 (97.5%)	1600 (97%)	154 (99.4%)
Time interval between birth of the 1st and 2nd twin, mean(sd)		30,746	13(22.2)	19.5 (736.7)	21 (361.7)	11.9 (17.3)

Abbreviations: LH, left-handed; RH, right-handed, MH, mixed-handed. Values are presented as mean and standard deviation (sd) or number (n) and frequency in the group (%).

definition (see **Supplementary Table 3**). The largest correlation coefficients using the left-handedness definition were with mother's handedness (r=.139). The largest correlation coefficients using the mixed-handedness definition were with sex (r=.191), monoamnionicity (r=.190), and neurodevelopmental delay (r=.173).

Test of associations: As explained above, we analyzed each variable individually with GEE logistic regression to correct for relatedness of twins (see **Table 2** and **Supplementary Table 4**). The higher probability of being LH, MH, or NRH was associated (at P<0.001) with being male (odd ratio (OR) 1.2, 95% confidence interval (CI) [1.14, 1.26] for left-handedness; OR 2.27, 95% CI [1.93, 2.67] for mixed-handedness; OR, 1.26, 95% CI[1.2, 1.32] for non-right-handedness), and shorter gestational age (OR 0.91, 95% CI [0.86, 0.95] for left-handedness; OR 0.78, 95% CI [0.67, 0.92] for mixed-handedness; OR 0.90, 95% CI[0.85, 0.94] for non-right-handedness). In preterm born twins, the prevalence of left-handed (16%) and mixed-handed (2%) were higher than in term born twins (14% and 1% respectively). The differences were even stronger in extreme preterm twins (20% left-handed and 5% mixed-handed) (see **Supplementary Table 5**).

Parental handedness was associated with handedness using the left-handedness definition and non-righthandedness definition (mother's handedness to left-handedness OR 1.60, 95% CI [1.48, 1.73] and to non-right-handedness OR 1.47, 95% CI [1.38, 1.58]; father's handedness to left-handedness OR 1.27, 95% CI [1.18, 1.37] and to non-right-handedness OR 1.23, 95% CI [1.15, 1.31]), but not significantly using mixed-handedness definition although effect sizes were similar. Breastfeeding was associated with handedness using the left-handedness (being breastfed to left-handedness OR 0.90, 95%CI [0.85, 0.95]). Mixed-handed twins had 3-fold higher risk of neurodevelopmental delay at 5 years (OR 2.51, 95% CI [1.75, 3.59]) and externalizing problems at 7 years old (OR 2.81, 95% CI [1.86, 4.22]).

Twin-specific characteristics. The left-handedness and mixed-handedness were equally common in MZ and DZ twins (15.3% vs 14.6%, $\chi^2(1)=3.4$, p=0.07 for left-handedness; and 1.2% vs 1.4%, $\chi^2(2)=1.4$, p=0.24 for mixed-handedness) and in 1st and 2nd born twins (15.2% vs 14.5%, $\chi^2(1)= 3.97$, p=0.05 for left-handedness; and 1.2% vs 1.4 % $\chi^2(1)= 2.8$, p=0.15 for mixed-handedness). We tested the hypothesis of the decreased prevalence in left-handedness in opposite-sex and same-sex female twins, and did not observe significant differences in any of three definitions of handedness (LH: $\chi^2(1)=5.57$, p=0.02; MH: $\chi^2(1)=0.58$, p=0.45; NRH: $\chi^2(1)=5.59$, p=0.018; see **Supplementary Table 6**). None of the twin-specific characteristics were associated with handedness.

Discussion

Our aim was to investigate the association of early life characteristics and handedness in a population-based sample of twins. The operational definition of handedness displays some variation in the literature, which may account for inconsistencies that we noted in our review of the literature. We therefore considered three definitions: right- versus left-handed, right- versus mixed-handed, and right- versus non-right-handed. Using the NTR data resources, we tested the association of handedness with a number of

twin-specific early life characteristics, which have not been considered before. These include amnionicity and time interval between birth of the 1st and 2nd twin. We also tested the association with bowel and bladder toilet skills delay at age 5 (indicators of neurodevelopment), and the association with aggression at 7 years old (a measure related to externalizing).

Our results support the higher incidence of left-handed individuals among twins (14.9%) in comparison with general population, which has also been reported in a meta-analysis of 59 twin studies (9.13% left-handed twins vs. 6.97% left-handed singletons)[56]. Consistent with the results of this meta-analysis, we found comparable prevalences in mixed-handedness among twin and singletons with 1.3% mixed-handed twins in our study vs. 2.67% mixed-handed singletons[56].

Parental handedness is associated with left-handedness, but not with mixed-handedness. This is consistent with reports on the association of left-handedness with mother's left-handedness and mixed-handedness20 and with fathers52. The absence of an association of parental handedness and mixed-handedness is consistent with results of Zhu et al. [17].

We confirm a higher prevalence of left-handedness and mixed-handedness in males and in preterm born individuals (<37 weeks and <28 weeks) in comparison with twins born after 37 week of gestation. In contrast to others (see **Table 1**), we found no evidence that twins' birthweight is associated with handedness. De Kovel et al. (2019) reported that left-handedness was associated with lower birthweight. Birthweight in twins is likely be partly affected by other factors than in singletons, as intrauterine growth restrictions play an important role in twin pregnancies[57].

Effects of season of birth or of sharing the womb with a male cotwin have been hypothesized as a result in variation in testosterone exposure. We found no effect of being born in summer months or of sharing the womb with a male cotwin. This is in contrast, with de Kovel (2018), who suggested that seasonal effects on handedness are present in females, possibly attributable to seasonal variation in maternal testosterone. Another disputed hypothesis is that females from opposite-sex twin pairs are exposed to higher testosterone levels during pregnancy that can affect lateralization and right-handedness. Vuoksimaa et al[14] compared 1,455 females from same-sex twin pares and 755 females from opposite-sex twin pairs, and found significantly lower prevalence of left-handedness in opposite-sex females (5.3%) compared to same-sex females (8.6%). However, our findings in 11,788 females from same-sex twin pairs and 6,000 females from opposite-sex twin pairs did not support those findings.

Interestingly, our findings confirm that breastfeeding is associated with right-handedness[15,51]. We can speculate on three potential pathways underlying this association. First, breastfeeding can be a confounding, either genetic or environmental, factor or a part of a causal pathway of some other risk factor. Caesarean section may adversely affect infant breastfeeding[58]. Insufficient lactation in mothers of multiple births can cause early breastfeeding cessation[59]. Second, a mother of a twin pair (or caregiver in case of bottle feeding) may tend to feed each twin in a certain position instead of switching feeding directions that can affect neurodevelopment. The discordant nipple preference and further opposite hand preference was described in several cases of

TABLE 3.
Univariate analysis of handedness and early life characteristics

	Left-handedness (RH vs LH) N = 36,997			Mixed-handedness (RH vs MH) *N* = 31,923			Non-right-handedness (RH vs LH+MH) *N* = 37,495		
	N	**OR (95%CI)**	***P***	***N***	**OR (95%CI)**	***P***	***N***	**OR (95%CI)**	***P***
Handedness-associated early life characteristics									
Sex (F/M)	36,997	1.2 (1.14 , 1.26)	$6.7x10^{-10}$	31,923	2.27 (1.93 , 2.67)	$1.4x10^{-16}$	37,495	1.26 (1.2 , 1.32)	$4.7x10^{-16}$
Gestational age, *continuous dichotomous (<37, >37 weeks)*	35,795	— 0.91 (0.86 , 0.95)	0.0002	30,883	— 0.78 (0.67 , 0.92)	$4.9x10^{-05}$	36,277	— 0.9 (0.85 , 0.94)	$2.6x10^{-06}$
Mother's handedness*	30,349	1.6 (1.48 , 1.73)	$1.1x10^{-23}$	26,179	1.67 (1.12 , 2.49)	0.034	31,869	1.4	$6.7x10^{-21}$
Father's handedness*	30,088	1.27 (1.18 , 1.37)	$3.1x10^{-07}$	25,922	1.97 (1.37 , 2.84)	0.002	31,643	7 (1.38 , 1.58)	$4.8x10^{-07}$
Breastfeeding (no/yes)	31,624	0.9 (0.85 , 0.95)	0.0010	27,261	0.99 (0.83 , 1.18)	0.953	32,038	1.23 (1.15 , 1.31)	0.0014
Aggression score at 7 years, *continuous dichotomous (< 18, > 18 points)*	20,327	— 1 (0.81 , 1.23)	0.244	17,516	— 2.81 (1.86 , 4.22)	$1.5x10^{-12}$	20,595	0.9 (0.86 , 0.95) — 1.14 (0.94 , 1.38)	0.004 —
Neurodevelopmental delay at 5 years (no/yes)	36,615	1.02 (0.87 , 1.2)	0.844	31,599	2.51 (1.75 , 3.59)	$2.4x10^{-05}$	37,106	1.14 (0.98 , 1.33)	0.143
Year of birth, *continuous dichotomous (< 2000, >2000)*	36,997	— 0.97 (0.92 , 1.02)	0.526	31,923	— 0.85 (0.72 , 1)	0.621	37,495	— 0.96 (0.91 , 1)	0.628
Mother's age at birth, *continuous dichotomous (<40, >40 years)*	35,885	— 0.87 (0.72 , 1.06)	0.199	30,955	— 1.18 (0.69 , 2)	0.912	36,364	— 0.9 (0.74 , 1.08)	0.204
Father's age at birth, *continuous dichotomous (<40, >40 years)*	35,520	— 0.92 (0.84 , 1)	0.454	30,621	— 1.4 (1.1 , 1.79)	0.453	35,990	— 0.95 (0.88 , 1.04)	0.626
Mode of conception (Natural/ Artificial)	34,058	0.95 (0.89 , 1.01)	0.133	29,380	1.03 (0.86 , 1.24)	0.764	34,526	0.95 (0.9 , 1.01)	0.173
Maternal smoking (no/yes)	35,878	1.08 (1.02 , 1.16)	0.039	30,945	1.33 (1.1, 1.61)	0.015	36,360	1.1 (1.04, 1.17)	0.009

Table 3 (continued)

	Left-handedness (RH vs LH) N = 36,997			Mixed-handedness (RH vs MH) *N* = 31,923			Non-right-handedness (RH vs LH+MH) *N* = 37,495		
	N	**OR (95%CI)**	***P***	***N***	**OR (95%CI)**	***P***	***N***	**OR (95%CI)**	***P***
Maternal stress during pregnancy (no/yes)	7,117	1.11 (0.99 , 1.25)	0.135	6,134	1.11 (0.77, 1.61)	0.642	7,202	1.11 (0.99, 1.24)	0.126
Being born in summer (no/yes)	36,993	1.06 (1 , 1.12)	0.102	31,919	1.08 (0.91, 1.29)	0.473	37,491	1.06 (1, 1.12)	0.083
Fetal presentation (cephalic/ non-cephalic)	7,182	0.95 (0.85 , 1.07)	0.479	6,189	1.48 (1.03, 2.13)	0.077	7,268	0.99 (0.88, 1.1)	0.826
Mode of delivery (vaginal/ instrumental)	31,373	0.97 (0.92 , 1.03)	0.426	27,090	1.03 (0.87, 1.21)	0.799	31,818	0.98 (0.93, 1.03)	0.510
Birthweight, *continuous dichotomous (<2500, >2500g)*	35,568	— 0.98 (0.94 , 1.03)	0.125	30,677	— 0.85 (0.72, 0.99)	0.002	36,045	— 0.97 (0.92, 1.02)	0.019
Apgar score 1 min, *continuous dichotomous (<7, >7 points)*	5,217	— 1.03 (0.84 , 1.26)	0.506	4,515	— 0.45 (0.28, 0.74)	0.025	5,281	— 0.95 (0.79, 1.15)	0.971
Twin-specific characteristics									
Birthorder (1st/2nd)	36,968	0.94 (0.9, 0.99)	0.042	31,896	1.17 (1.01, 1.34)	0.077	37,465	0.96 (0.92, 1)	0.138
Zygosity (DZ/MZ)	36,827	0.95 (0.9, 1)	0.076	31,777	1.12 (0.95, 1.33)	0.261	37,321	0.96 (0.91, 1.01)	0.169
Chorionicity (DC/MC)	11,043	1.06 (0.96, 1.16)	0.347	9,544	0.81 (0.59, 1.13)	0.304	11,201	1.03 (0.94, 1.13)	0.567
Amnionicity (DA/MA)	11,000	1.21 (0.92, 1.6)	0.260	9,506	0.24 (0.05, 1.24)	0.153	11,155	1.13 (0.86, 1.48)	0.479
Time between birth of 1st and 2nd twin, *continuous dichotomous (<30 min, >30 min)*	30,312	— 0.99 (0.91, 1.08)	0.113	26,167	— 0.89 (0.68, 1.16)	0.368	30,752	— 0.98 (0.9, 1.06)	0.213

RH, right-handed; LH, left-handed; MH, mixed-handed; NRH, non-right-handed (LH+MH); MZ, monozygotic; DZ, dizygotic; MC, monochorionic; DC, dichorionic; MA, monoamniotic; DA, diamniotic; *P, p*-value in regression with correction for relatedness (GEE). Blue, early life factor significant for all three definitions of handedness. Green, early life factor significant for LH and NRH. Yellow, early life factor significant for MH. Level of significance α = 0.001109.
*Mother's and father's handedness are included in the regression analysis in the same definition as the child handedness.

twins in studies of primates[60]. Finally, in comparison to bottle feeding, breastfeeding can increase myelinization and neurodevelopment[51,61]. Further investigations are needed to analyse these hypotheses.

A novel finding of our study is the relationship between mixed-handedness and later outcomes. Mixed-handed twins had a higher risk of problematic toilet skill training (relevant to neurodevelopment) at 5 years and aggression (relevant to externalizing) at 7 years. Mental health problems tend to be more severe in non-right-handed adolescents[30]. According to our findings, one of the phenotypic markers associated with increased risk to adverse neurodevelopment could be the mixed-handedness, but not left-handedness. We found no association of left-handedness with externalizing problems, while some previous studies found a higher rate of left-handedness in clinical samples of children with psychiatric disorders such as oppositional defiant disorder[29].

A limitiation of the current study is that the measurement of handedness was based on parental report based on a single item. We found that this item captures most of the other information in questions on hand used by child while drinking from the cup, eating, throwing a ball, picking up the coin, and combing hair. For some participants data on early life characteristics was missing, due to the absence of the relevant questions in earlier surveys (e.g. fetal presentation at birth, maternal stress during pregnancy). The data on chorionicity were obtained from the Dutch national pathology database, after linking to the NTR data, but did not include all twins.

In conclusion, associations of handedness based on three definitions with a wide range of early life characteristics in a national twin cohort depend on the definition of handedness. Sex and gestational age are associated with all definitions of handedness, but there is some variation in associations with other early life characteristics and left- or mixed-handedness. Left- and non-right-handedness were associated with parental left-handedness and mixed-handedness was associated with higher risk of neurodevelopmental delay at 5 years old and externalizing problems at 7 years old. Twin-specific characteristics were not related to handedness.

References

1. Hepper, P. G. The developmental origins of laterality: Fetal handedness. *Developmental Psychobiology* **55**, 588–595 (2013).
2. Hepper, P. G., Shahidullah, S. & White, R. Handedness in the human fetus. *Neuropsychologia* **29**, 1107–1111 (1991).
3. Parma, V., Brasselet, R., Zoia, S., Bulgheroni, M. & Castiello, U. The origin of human handedness and its role in pre-birth motor control. *Scientific Reports* **7**, 16804 (2017).
4. Medland, S. E. *et al.* Genetic influences on handedness: data from 25,732 Australian and Dutch twin families. *Neuropsychologia* **47**, 330–337 (2009).
5. Annett, M. *Left, right, hand, and brain: the right shift theory*. (L. Erlbaum Associates, 1985).
6. McManus, I. Handedness, language dominance and aphasia: a genetic model. Psychological Medicine. *Monograph Supplement* **8**, 3–40 (1985).
7. McManus, C. Is any but a tiny fraction of handedness variance likely to be due to the external environment? *Laterality* 1–5 (2021) doi:10.1080/1357650X.2021.1892126.
8. Odintsova, V. *et al.* DNA methylation in peripheral tissues and left-handedness. *Scientific Reports,* 12(1), 5606 (2022).
9. Bakan, P., Dibb, G. & Reed, P. Handedness and birth stress. *Neuropsychologia* **11**, 363–366 (1973).
10. Geschwind, N. & Behan, P. Left-handedness: association with immune disease, migraine, and developmental learning disorder. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 5097–5100 (1982).
11. Geschwind, N. & Galaburda, A. M. Cerebral lateralization. Biological mechanisms, associations, and pathology: I. A hypothesis and a program for research. *Archives of neurology* **42**, 428–459 (1985).
12. Grimshaw, G. M., Bryden, M. P. & Finegan, J.-A. K. Relations Between Prenatal Testosterone and Cerebral Lateralization in Children. in (1995).
13. Gadea, M., Gómez, C., González-Bono, E., Salvador, A. & Espert, R. Salivary testosterone is related to both handedness and degree of linguistic lateralization in normal women. *Psychoneuroendocrinology* **28**, 274–287 (2003).
14. Vuoksimaa, E., Eriksson, C. J. P., Pulkkinen, L., Rose, R. J. & Kaprio, J. Decreased prevalence of left-handedness among females with male co-twins: evidence suggesting prenatal testosterone transfer in humans? *Psychoneuroendocrinology* **35**, 1462–1472 (2010).
15. De Kovel, C. G. F., Carrión-Castillo, A. & Francks, C. A large-scale population study of early life factors influencing left-handedness. *Scientific reports* **9**, 584 (2019).
16. Searleman, A., Porac, C. & Coren, S. Relationship between birth order, birth stress, and lateral preferences: a critical review. *Psychological bulletin* **105**, 397–408 (1989).
17. Zhu, J. L. *et al.* Infertility, infertility treatment, and mixed-handedness in children. *Early human development* **85**, 745–749 (2009).
18. Dragović, M., Milenković, S., Kocijancić, D. & Zlatko, S. Etiological aspect of left-handedness in adolescents. *Srpski arhiv za celokupno lekarstvo* **141**, 354–358 (2013).
19. Bailey, L. M. & McKeever, W. F. A large-scale study of handedness and pregnancy/birth risk events: implications for genetic theories of handedness. *Laterality* **9**, 175–188 (2004).
20. Johnston, D. W., Nicholls, M. E. R., Shah, M. & Shields, M. A. Handedness, health and cognitive development: evidence from children in the National Longitudinal Survey of Youth. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **176**, 841–860 (2013).
21. Karev, G. B. Season of birth and parental age in right, mixed and left handers. *Cortex; a journal devoted to the study of the nervous system and behavior* **44**, 79–81 (2008).
22. Gutteling, B. M., de Weerth, C. & Buitelaar, J. K. Prenatal Stress and Mixed-Handedness. *Pediatric Research* **62**, 586–590 (2007).
23. Obel, C., Hedegaard, M., Henriksen, T. B., Secher, N. J. & Olsen, J. Psychological factors in pregnancy and mixed-handedness in the offspring. *Developmental medicine and child neurology* **45**, 557–561 (2003).

24. Domellöf, E., Johansson, A.-M. & Rönnqvist, L. Handedness in preterm born children: a systematic review and a meta-analysis. *Neuropsychologia* **49**, 2299–2310 (2011).
25. Van Heerwaarde, A. A. *et al.* Non-right-handedness in children born extremely preterm: Relation to early neuroimaging and long-term neurodevelopment. *PloS one* **15**, e0235311 (2020).
26. Darvik, M., Lorås, H. & Pedersen, A. V. The Prevalence of Left-Handedness Is Higher Among Individuals With Developmental Coordination Disorder Than in the General Population. *Frontiers in psychology* **9**, 1948 (2018).
27. Dinsdale, N. L., Reddon, A. R. & Hurd, P. L. Sex differences in the relationship between aggressiveness and the strength of handedness in humans. *Laterality* **16**, 385–400 (2011).
28. Van der Feen, F. E., Zickert, N., Groothuis, T. G. G. & Geuze, R. H. Does hand skill asymmetry relate to creativity, developmental and health issues and aggression as markers of fitness? *Laterality* **25**, 53–86 (2020).
29. Logue, D. D., Logue, R. T., Kaufmann, W. E. & Belcher, H. M. E. Psychiatric disorders and left-handedness in children living in an urban environment. *Laterality* **20**, 249–256 (2015).
30. Van der Hoorn, A. *et al.* Non-right-handedness and mental health problems among adolescents from the general population: The Trails Study. *Laterality* **15**, 304–316 (2010).
31. Davis, A. & Annett, M. Handedness as a function of twinning, age and sex. *Cortex; a journal devoted to the study of the nervous system and behavior* **30**, 105–111 (1994).
32. Ellis, S. J., Ellis, P. J. & Marshall, E. Hand preference in a normal population. *Cortex; a journal devoted to the study of the nervous system and behavior* **24**, 157–163 (1988).
33. Heikkilä, K. *et al.* Higher Prevalence of Left-Handedness in Twins? Not After Controlling Birth Time Confounders. *Twin Research and Human Genetics* **18**, 526–532 (2015).
34. Heikkilä, K. *et al.* Triplets, birthweight, and handedness. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 6076–6081 (2018).
35. Sicotte, N. L., Woods, R. P. & Mazziotta, J. C. Handedness in Twins: A Meta-analysis. *Laterality: Asymmetries of Body, Brain and Cognition* **4**, 265–286 (1999).
36. Vuoksimaa, E., Koskenvuo, M., Rose, R. J. & Kaprio, J. Origins of handedness: a nationwide study of 30,161 adults. *Neuropsychologia* **47**, 1294–1301 (2009).
37. Derom, C., Thiery, E., Vlietinck, R., Loos, R. & Derom, R. *Handedness in Twins According to Zygosity and Chorion Type: A Preliminary Report. Behavior Genetics* vol. 26 https://link.springer.com/content/pdf/10.1007/BF02359484.pdf (1996).
38. Medland, S. E. *et al.* Special Twin Environments, Genetic Influences and their Effects on the Handedness of Twins and their Siblings. *Twin Research* **6**, 119–130 (2003).
39. Van Beijsterveldt, C. E. M. *et al.* Chorionicity and Heritability Estimates from Twin Studies: The Prenatal Environment of Twins and Their Resemblance Across a Large Number of Traits. *Behavior genetics* **46**, 304–314 (2016).
40. James, W. H. & Orlebeke, J. F. Determinants of handedness in twins. *Laterality* **7**, 301–307 (2002).
41. Willemsen, G., Odintsova, V., de Geus, E. & Boomsma, D. I. Twin-singleton comparisons across multiple domains of life. in *Twin and Higher-order Pregnancies* (eds. Khalil, A., Lewi, L. & Lopriore, E.) (Springer International Publishing, 2021).
42. Ligthart, L. *et al.* The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-Family Designs. *Twin Research and Human Genetics* **22**, 623–636 (2019).
43. Francis, K., Mannion, A. & Leader, G. The Assessment and Treatment of Toileting Difficulties in Individuals with Autism Spectrum Disorder and Other Developmental Disabilities. *Review Journal of Autism and Developmental Disorders* **4**, 190–204 (2017).
44. Schum, T. R. *et al.* Sequential Acquisition of Toilet-Training Skills: A Descriptive Study of Gender and Age Differences in Normal Children. *Pediatrics* **109**, e48–e48 (2002).
45. Achenbach, T. M. & Rescorla, L. *Manual for the ASEBA school-age forms & profiles : an integrated system of multi-informant assessment.* (2001).

46. Orlebeke, J. F., Knol, D. L., Koopmans, J. R., Boomsma, D. I. & Bleker, O. P. Left-handedness in twins: genes or environment? *Cortex; a journal devoted to the study of the nervous system and behavior* **32**, 479–490 (1996).
47. Sutcliffe, A. G. *et al.* Laterality in five-year-olds conceived by intracytoplasmic sperm injection, standard in vitro fertilisation and natural conception: a European study. *BJOG : an international journal of obstetrics and gynaecology* **112**, 1397–1401 (2005).
48. Ooki, S. Nongenetic factors associated with human handedness and footedness in Japanese twin children. *Environmental health and preventive medicine* **11**, 304–312 (2006).
49. Denny, K. Breastfeeding predicts handedness. *Laterality* **17**, 361–368 (2012).
50. Van der Elst, W. *et al.* On the association between lateral preferences and pregnancy/birth stress events in a nonclinical sample of school-aged children. *Journal of clinical and experimental neuropsychology* **33**, 1–8 (2011).
51. Hujoel, P. P. Breastfeeding and handedness: a systematic review and meta-analysis of individual participant data. *Laterality* **24**, 582–599 (2019).
52. Fagard, J., de Agostini, M., Huet, V., Granjon, L. & Heude, B. Is Handedness at Five Associated with Prenatal Factors? *International journal of environmental research and public health* **18**, (2021).
53. Elkadi, S., Nicholls, M. E. & Clode, D. Handedness in opposite and same-sex dizygotic twins: testing the testosterone hypothesis. *Neuroreport* **10**, 333–336 (1999).
54. Marlow, N., Hennessy, E. M., Bracewell, M. A. & Wolke, D. Motor and executive function at 6 years of age after extremely preterm birth. *Pediatrics* **120**, 793–804 (2007).
55. Papadatou-Pastou, M. *et al.* Human handedness: A meta-analysis. *Psychological bulletin* **146**, 481–524 (2020).
56. Pfeifer, L. S. et al. Handedness in twins: meta-analyses. *BMC Psychology,* 10(1), 11 (2022).
57. Muhlhausler, B. S., Hancock, S. N., Bloomfield, F. H. & Harding, R. Are Twins Growth Restricted? *Pediatric Research* **70**, 117–122 (2011).
58. Hobbs, A. J., Mannion, C. A., McDonald, S. W., Brockway, M. & Tough, S. C. The impact of caesarean section on breastfeeding initiation, duration and difficulties in the first four months postpartum. *BMC Pregnancy and Childbirth* **16**, 90 (2016).
59. Damato, E. G., Dowling, D. A., Standing, T. S. & Schuster, S. D. Explanation for cessation of breastfeeding in mothers of twins. *Journal of human lactation : official journal of International Lactation Consultant Association* **21**, 296–304 (2005).
60. Hopkins, W. D. Laterality in Maternal Cradling and Infant Positional Biases: Implications for the Development and Evolution of Hand Preferences in Nonhuman Primates. *International journal of primatology* **25**, 1243–1265 (2004).
61. Deoni, S. C. L. *et al.* Breastfeeding and early white matter development: A cross-sectional study. *NeuroImage* **82**, 77–86 (2013).

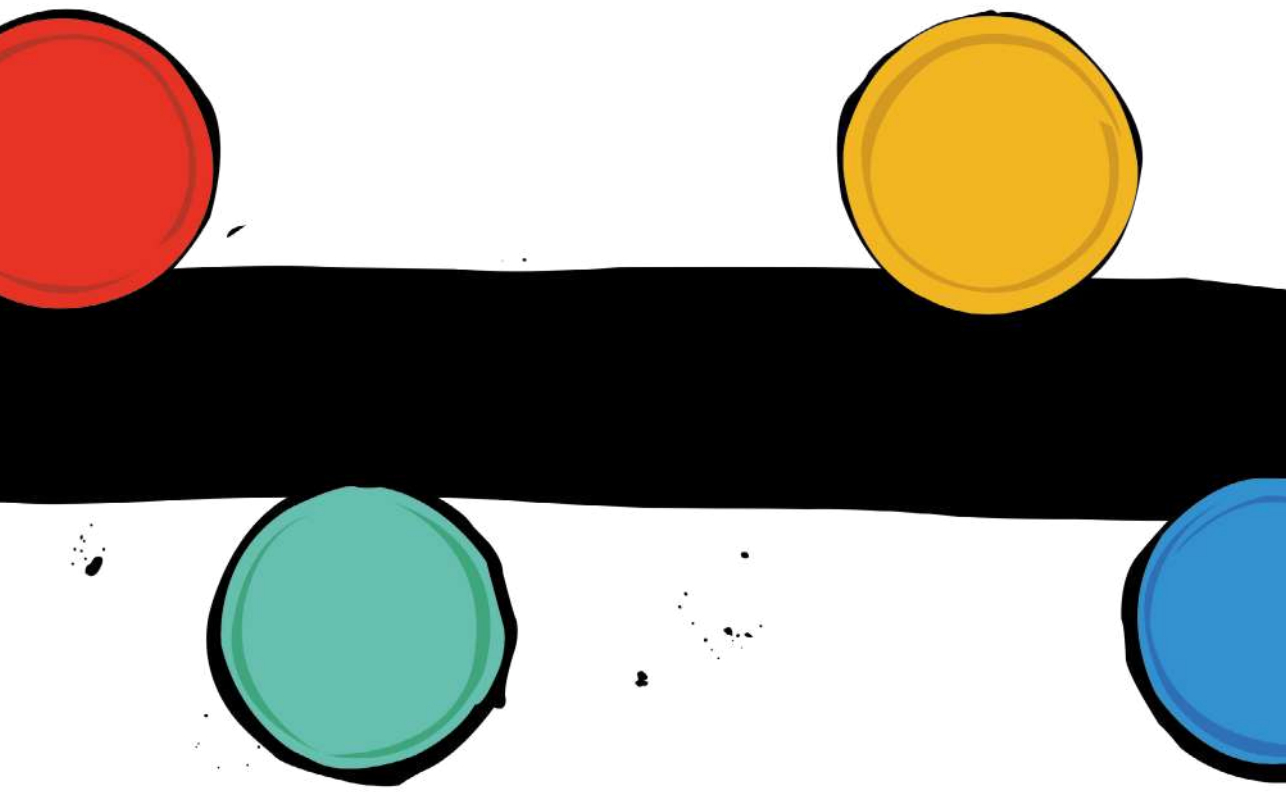
Supplements

The supplementary materials are available online at

https://surfdrive.surf.nl/files/index.php/s/sI202TIzClCZmAG

Chapter 6

DNA methylation in peripheral tissues and left-handedness



Abstract

Handedness has low heritability and epigenetic mechanisms have been proposed as an etiological mechanism. To examine this hypothesis, we performed an epigenome-wide association study (EWAS) of left-handedness. In a meta-analysis of 3,914 adults of whole-blood DNA methylation, we observed that CpG sites located in proximity of handedness-associated genetic variants were more strongly associated with left-handedness than other CpG sites (P=0.04), but did not identify any differentially methylated positions. We identified differentially methylated regions at 20q11.23 (P=0.00004) and 2p25.1 (P=0.03), which were less methylated in left-handed adults. In longitudinal analyses of DNA methylation in peripheral blood and buccal cells from children (N=1,737), we observed moderately stable associations across age (correlation range [0.355-0.578]) but inconsistent across tissues (correlation range [-0.384-0.318]). We conclude that DNA methylation in peripheral tissues captures little of the variance in handedness. Future investigations should consider other more targeted sources of tissue, such as the brain.

Published as: Odintsova V., Suderman M., Hagenbeek F., Caramaschi D., Hottenga J.J., Pool R., BIOS consortium, Dolan C.V., Ligthart L., van Beijsterveldt C.E.M., Willemsen G., de Geus E.J.C., Beck J.J., Ehli E.A., Cuellar-Partida G., Evans D.M., Medland S.E., Relton C.L., Boomsma D.I., van Dongen J. (2022) DNA methylation in peripheral tissues and left-handedness. *Scientific Reports,* 12(1): 5606.

Introduction

Handedness, defined as the preferential use of one hand over the other, is established early in life and represents a highly stable trait that is thought to be accompanied by changes in brain[1], corticospinal tract[2], peripheral innervation and vascularization of arm skeletal muscles[3], arm dynamics[4], and possibly the immune system[5]. Laterality is already observable in very early stages of development: fetuses show coordinated hand movements at 8–12 weeks post-conception with more right than left arm movements in 85% of fetuses[6–8]. In children and adults, the prevalence of left-handedness is about 10%[9]. Handedness clearly clusters in families, but its inheritance pattern is not clear and the heritability of handedness is relatively low: approximately 25% in twin studies (with 95% confidence intervals (CI) ranging from 11 to 30%[10–12]) and 11.9% (95% CI: 7.2–17.7) based on autosomal Identity By Descent (IBD) information from closely related individuals in the UK Biobank[13]. Early genetic hypotheses on the development of hand preference incorporated a component of randomness[14,15]: depending on which alleles were inherited, a person would be right-handed or have an equal chance of being either left- or right-handed. This randomness has also been referred to as "developmental instability", or "fluctuating asymmetry", representing developmental variance unique to the person[16]. Such randomness could explain monozygotic twin discordance in handedness[15] as reported in some twin studies[17–19], although very early studies did not confirm zygosity by DNA testing.

Candidate genes associated with handedness and brain and spinal asymmetry include leucine rich repeat transmembrane neuronal 1 *(LRRTM1)*[20], LIM domain only 4 *(LMO4)*[21], neuronal differentiation 6 *(NEUROD6)*[21], proprotein convertase subtilisin/kexin type 6 *(PCSK6)*[22,23] and the androgen receptor gene *(AR)*[24–26]. However, genome-wide association studies (GWASs) found no support for these candidate genes[27–30], but identified multiple novel loci. Recently, the largest genome-wide association study to date, which included more than 1,5 million right-handed and 194,000 left-handed individuals, found 41 loci associated with left-handedness[13]. Besides previously described associations of left-handedness with loci that contain microtubule-associated protein 2 *(MAP2)*[29,30] and tubulin beta class 1 *(TUBB)*[29], the list of genome-wide significant associations was expanded with other microtubule formation and regulation genes *(TUBB3, NDRG1, TUBB4A, TUBA1B, BUB3 and TTC28)*[13]. Thus, multiple variants were close to genes involved in microtubule functions that form part of the cytoskeleton, and play a role in neurogenesis, axon transport[31], and brain asymmetry[32]. The results of functional analyses suggested involvement of neurogenesis and the central nervous system and brain tissues, including hippocampus and cerebrum, in the etiology of left-handedness. The variance of handedness explained by single nucleotide polymorphisms (SNP heritability) on the liability scale was 3.45% in this meta-analysis[13]. The estimate in UK Biobank was 5.9%[13].

Partly because of the limited success of early genetic association studies, epigenomic studies have been proposed as promising targets to identify mechanisms underlying handedness[33–35]. Epigenome-wide association studies (EWAS) perform association tests for several hundred thousand of CpGs (cytosine-phosphate-guanine nucleotide base pairing) to identify differentially methylated positions (DMPs) associated with a trait.

Approaches that test associations across multiple nearby correlated CpGs to identify differentially methylated regions (DMRs)[36], or that combine multiple CpGs into DNA methylation scores[37] help improve power by combining the effects of multiple CpG sites and reducing the number of conducted tests. The predictive value of DNA methylation by construction of individual methylation scores has been shown for several outcomes, e.g. body mass index[38]. Epigenetic variation could be one pathway to connect the hypothesized random component of handedness, and contribute to asymmetrical gene expression in the two brain hemispheres[39] and the spinal cord[40]. The latter was supported by a genome-wide DNA methylation analysis in the right and left part of the fetal spinal cord from six samples obtained between 8 and 12 weeks post conception that detected asymmetrically methylated CpG islands of several genes[40].

At present, no epigenome-wide association studies of handedness have been performed, and the role of DNA methylation in handedness has only been examined in small candidate-gene studies[41,42]. Here we analyzed DNA methylation data and left-handedness from two cohorts – the Netherlands Twin Register (NTR) and Avon Longitudinal Study of Parents and Children (ALSPAC). Both cohorts include methylation data in children and adults. In the ALSPAC cohort, adults and children are related (parents and offspring), and in the NTR cohort adults and children come from independent samples. We excluded ambidextrous and mixed-handed persons, and treated handedness as a dichotomous trait (left- or right-handed). First, we performed a meta-analysis of DNA methylation data from the two largest groups with DNA methylation data in adults (total sample size=3,914) to identify differentially methylated positions and differentially methylated regions associated with left-handedness. Next, we performed additional analyses in which we 1) examined if the epigenetic signal for left-handedness was enriched near previously reported GWAS loci[13]; 2) examined methylation differences between left- and right-handed twins from discordant monozygotic (MZ) twin pairs; 3) characterized the longitudinal and cross-tissue similarity of the genome-wide epigenetic signal associated with left-handedness using data from children; and 4) created methylation scores and estimated their predictive value over and above polygenic scores (**Figure 1**).

Results

Epigenome-wide association meta-analysis of left-handedness

Tables 1-2 and **Supplementary Tables 1-4** display the characteristics of the participants included in the study. The epigenome-wide association study of left-handedness meta-analysed data from adults (*N*=3,914) with DNA methylation data in peripheral blood (Illumina, 450k) from NTR (*N*=2,682, 34% male, mean age at methylation 36.5, standard deviation (SD) 12.7) and ALSPAC (*N*=1,232, 30% male, mean age at methylation 48.98, SD 5.55). In EWAS discovery cohorts, the prevalence of left-handedness was 12% in NTR, and 8% in ALSPAC. The prevalence of left-handedness as a function of year of birth in NTR is provided in **Supplementary Table 2**.

We tested 409,562 CpGs with adjustment for age, sex, smoking status, body mass index (BMI), measured or estimated cell proportions, and technical covariates. Genome-wide

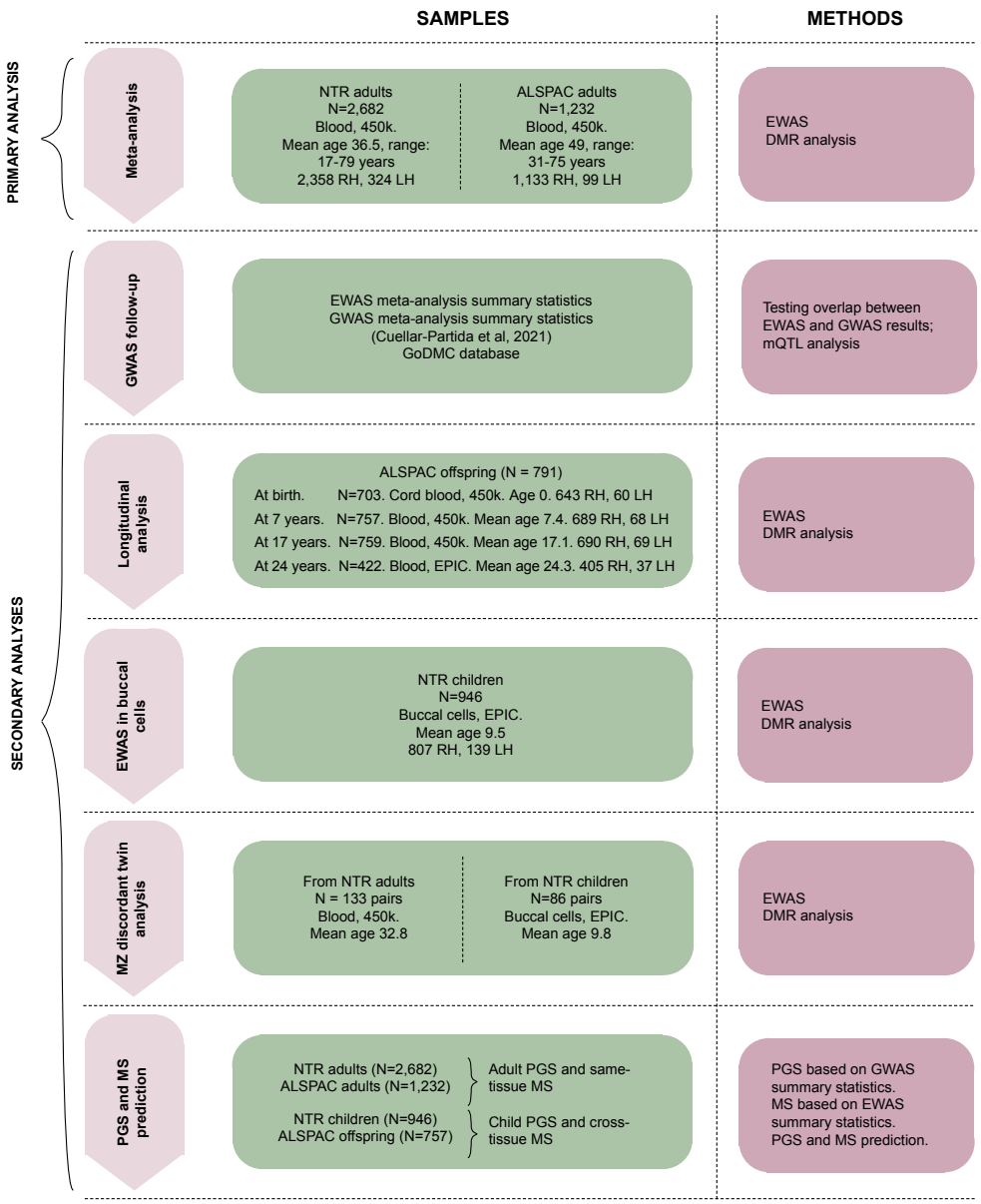


Figure 1: Flowchart of epigenome-wide association study of left-handedness

The flowchart summarizes the study design. The primary analyses included EWASs in NTR adults and ALSPAC adults, followed by meta-analysis to identify DNA methylation sites associated with left-handedness. The secondary analyses included: 1) left-handedness GWAS loci follow-up; 2) longitudinal analysis of DNA methylation at four ages in ALSPAC offspring; 3) analysis of buccal cell DNA methylation in NTR children; 4) analysis of DNA methylation differences between left and right-handed co-twins from NTR discordant MZ twin pairs and 5) polygenic and DNA methylation scores prediction. For left-handedness prediction, polygenic scores (PGS) were created based on left-handedness GWAS summary statistics not including NTR/ALSPAC. Methylation scores (MS) were created based on weights from EWASs in NTR adults, ALSPAC adults, NTR children and ALSPAC offspring at 7 years old to estimate the predictive performance.

LH, left-handed; RH, right-handed. DMR, differentially methylated region. EWAS, epigenome-wide association study. GWAS, genome-wide association study. GoDMC, Genetics of DNA Methylation Consortium. mQTL, methylation quantitative trait locus. Blood, buccal cells indicate tissue of DNA methylation. 450k, EPIC indicate the platform for DNA methylation measurement.

TABLE 1.
Characteristics of adult cohorts included in the primary meta-analysis

	NTR adults *N* = 2682		**ALSPAC adults *N* = 1232**	
	LH	**RH**	**LH**	**RH**
N	324 (12%)	2358 (88%)	99 (8%)	1133 (92%)
Age at blood sampling	34.3 (11.2)	36.8 (12.9)	49.1 (5.9)	49.0 (5.5)
Sex				
Males	119 (37%)	783 (33%)	31 (31%)	333 (29%)
Females	205 (63%)	1575 (67%)	68 (69%)	800 (71%)
Multiples	315 (97.2%)	2171 (92%)	0	0
BMI	24.2 (3.7)	24.2 (3.9)	25.80 (4.3)	26.71 (4.7)
Smoking (current)	65 (20%)	486 (21%)	37 (37%)	424 (37%)
Cell proportion	Neutrophils		B lymphocytes	
	52.8 (8.7)	52.4 (9.2)	10.8 (4.3)	10.4 (4)
	Eosinophils		CD4T	
	3.1 (1.9)	3.1 (2.3)	18.4 (7.2)	18 (6.6)
	Monocytes		CD8T	
	8.4 (2.2)	8.4 (2.4)	2.2 (3.8)	1.9 (3.1)
			Natural killer cells	
			21.9 (6.5)	20.7 (5.7)
			Granulocytes	
			45.5 (14.7)	47.9 (12.4)
			Monocytes	
			7.5 (3.4)	7.4 (3.5)

NTR, Netherlands Twin Register. ALSPAC, Avon Longitudinal Study of Parents and Children.
LH, left-handed. RH, right-handed.
Whole blood DNA methylation (Illumina 450k). Numbers in EWAS basic models are reported
Values are presented as mean (standard deviation (SD)) or n (%).
Current smokers in ALSPAC were defined as those with cg05575921 methylation below 0.82 (see Methods).

TABLE 2.
Characteristics of the datasets included in secondary analyses

	ALSPAC offspring (longitudinal)								NTR children	
	at birth *N*=703		7 years old *N*=757		17 years old *N*=759		24 years old *N*=442		*N*=946	
	LH	**RH**	**LH**	**RH**	**LH**	**RH**	**LH**	**RH**	**LH**	**RH**
N	60 (8.5%)	643 (91.5%)	68 (9%)	689 (91%)	69 (9%)	690 (91%)	37 (8.4%)	405 (91.6%)	139(15%)	807(85%)
Age at blood sampling			7.43 (0.1)	7.4 (0.1)	16.9 (1.1)	17.1 (1.0)	24.3 (0.7)	24.3 (0.7)	9.58(1.78)	9.56(1.86)
Sex										
Males	32 (53.3%)	302 (47%)	36 (52.9%)	332 (48.2%)	36 (52.2%)	321 (46.5%)	16 (43.2%)	173 (42.7%)	71 (51%)	412 (51%)
Females	28 (46.7%)	341 (53%)	32 (47.1%)	357 (51.8%)	33 (47.8%)	369 (53.5%)	21 (56.8%)	232 (57.3%)	68 (49%)	395 (49%)
Gestational age	39.6 (1.5)	39.6 (1.4)	39.6 (1.4)	39.6 (1.6)					35.51(2.83)	35.93(2.52)
Birth weight	3567.8 (434.4)	3477.3 (493.6)	3583 (445.4)	3481 (493.3)					2369 (585.2)	2407 (533.5)
Maternal smoking during pregnancy	9 (15.2%)	74 (11.6%)	10 (14.9%)	81 (11.8%)					14(11%)	56(7%)
BMI					22.5 (3.6)	22.5 (3.6)	24.2 (3.6)	24.4 (4.5)		
Smoking (current)					20 (29.4%)	161 (23.6%)	9 (24.3%)	131 (33%)		

NTR, Netherlands Twin Register. ALSPAC, Avon Longitudinal Study of Parents and Children. NTR: buccal cell DNA methylation. ALSPAC: cord blood and whole blood DNA methylation. Numbers in EWAS basic models are reported
LH, left-handed; RH, right-handed; BMI, body mass index
Values are presented as mean (SD) or n (%).
Current smokers in ALSPAC were defined as those with cg05575921 methylation below 0.82 (see Methods).

TABLE 3.
Top differentially methylated positions from EWAS meta-analysis of left-handedness

CpG	CHR	Position [a]	Gene	Location	β	SE	*P*	FDR	Direction
cg22804475	8	20054597	*ATP6V1B2*	TSS200	-0.0007	0.0002	$1.28x10^{-06}$	0.197	--
cg09239756	12	106642360	*CKAP4*	TSS1500	-0.0032	0.0007	$1.82x10^{-06}$	0.197	--
cg22541911	12	51785465	*GALNT6*	TSS1500	-0.0010	0.0002	$1.90x10^{-06}$	0.197	--
cg13719901	3	46608139	*LRRC2*	5'UTR; TSS200	0.0081	0.0017	$2.60x10^{-06}$	0.197	++
cg02850812	13	46961666	*C13orf18*	TSS200	-0.0014	0.0003	$2.62x10^{-06}$	0.197	--
cg16852837	18	51750955	*MBD2*	1st Exon; 5'UTR	-0.0006	0.0001	$3.28x10^{-06}$	0.205	--
cg09893588	20	61340109	*NTSR1*	TSS200	-0.0011	0.0003	$9.12x10^{-06}$	0.256	--
cg12402132	12	121148554	*UNC119B*	Gene body	-0.0009	0.0002	$9.54x10^{-06}$	0.256	--

β is the regression coefficient for left-handedness in EWAS meta-analysis adjusted model that included $N_{NTR\ adults}$ = 2,663 and $N_{ALSPAC\ adults}$ = 1,058. CpGs with uncorrected p-value < 1.0×10^{-5} are presented.
LH, left-handed; RH, right-handed; CHR, chromosome; SE, standard error; FDR, false discovery rate; TSS200, 200 base pairs upstream of transcription start site; TSS1500, 1500 base pairs upstream of transcription start site; 5'UTR, 5' untranslated region;
++ positive direction of effect in each cohort; --, negative direction of effect in each cohort
a Genome build Hg19 (build 37). See additional information on CpGs in **Supplementary Table 12**.

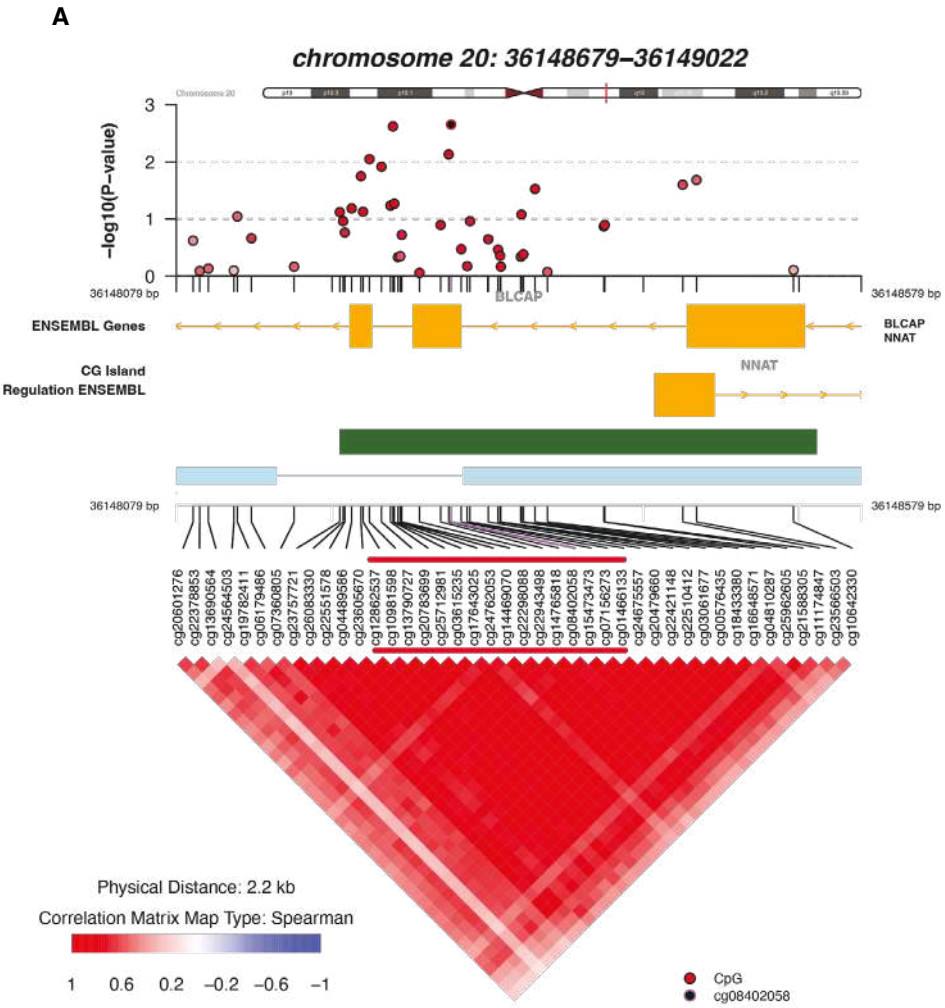
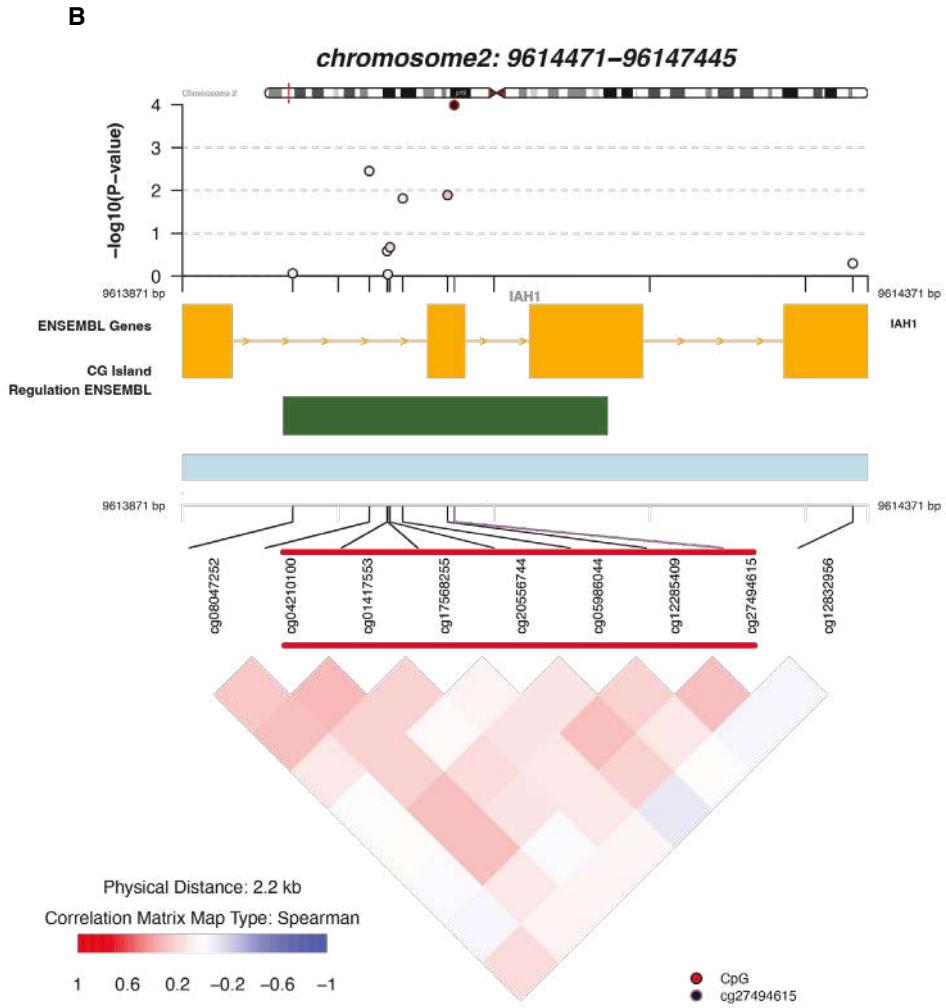


Figure 2. Differentially methylated regions associated with left-handedness in meta-analysis

The figure represents the differentially methylated regions at p-adjusted <0.05 in meta-analysis of DMR statistics across groups of NTR adults and ALSPAC adults (*N*= 3,712). The top panel of each plot depicts the EWAS *P*-values for CpGs in the differentially methylated regions **A)** at chromosome 20; and **B)** at chromosome 2. The x-axis indicates the position in base pair (bp) for the region, while y-axis indicates the strength of association from meta-analysis EWAS with left-handedness (adjusted model). The middle panel shows the genomic coordinates (genome build GRCh37/hg19) and the functional annotation of the region: the ENSEMBL Genes track shows the genes in the genomic region (orange); the CpG Island track shows the location of CpG islands (green); the Regulation ENSEMBL track shows regulatory



regions (blue). CpGs from DMR associated with left-handedness are indicated with red lines above the correlation heatmap. More detailed information on the regions is provided in Table 4. The bottom panel shows the Spearman correlation between methylation levels of CpGs in the window. The figure was computed using the R package CoMET. See additional information on CpGs from the regions in **Supplementary Table 12**.

EWAS test statistics from each cohort separately, and from the meta-analysis, showed no inflation (**Supplementary Tables 5-11**). None of the associations with CpG sites reached epigenome-wide significance (i.e. Bonferroni adjusted P<0.05 or false discovery rate < 5%) (**Supplementary Figure 1**). The CpGs with lowest p-values in meta-analysis ($P<1x10^{-5}$) are shown in **Table 3**. Six of eight CpGs were located near transcription start sites on different chromosomes: in the *LRRC2* gene on chromosome 3, in the *ATP6V1B2* gene on chromosome 8, in the *CKAP4, GALNTR6,* and *UNC1198* genes on chromosome 12, in the *C13orf18* gene on chromosome 13, in the *MBD2* gene on chromosome 18, and in the *NTSR1* gene on chromosome 20. The average difference in DNA methylation between left-handers and right-handers at these CpGs was small (from 0.06% to 0.8%; i.e. 0.0006 to 0.008 on the methylation beta-value scale) with lower methylation level in left-handers at all CpGs except for cg13719901 (LRRC2) (**Supplementary Figure 2** and **3**).

The DMR meta-analysis detected two DMRs associated with left-handedness (Fig.2, Table 4, Supplementary Table 12). The DMR on chromosome 20 (*BLCAP; NNAT;* 16 CpGs) had lower DNA methylation in left-handers than in right-handers (*P*-value adjusted for multiple testing (P_{adj}) =0.00004). The DMR on chromosome 2 (*IAH1;* 7 CpGs) also had lower DNA methylation in left-handers (P_{adj}=0.03). In total, 15 of 16 CpGs in the DMR on chromosome 20 (**Supplementary Figure 4** and **5**) and 6 of 7 CpGs in the DMR on chromosome 2 (**Supplementary Figure 6** and **7**) were hypomethylated in left-handers. The average absolute DNA methylation difference at these regions between left-handers and right-handers based on meta-analysis regression coefficients for the individual CpGs was 0.4% for the DMR on chromosome 20, and 0.1% for the DMR on chromosome 2. Both DMRs were within CpG islands, and were not detected in the individual cohorts. Some CpG sites within left-handedness-associated DMRs have been previously associated with other traits, these associations are listed in **Supplementary Table 13**.

GWAS follow-up

We tested the overlap of our EWAS meta-analysis results with findings from the most recent GWAS meta-analysis of handedness[13]. CpGs located within 1 Mb window of SNPs associated with left-handedness (at $P<5x10^{-8}$) were on average more strongly associated with left-handedness in the EWAS meta-analysis than the other tested CpGs (ß=0.027, *P*=0.04). The effect was weaker when less stringent GWAS p-value cut-offs were applied (i.e. SNPs with $P<1x10^{-6}$, and SNPs with $P<1x10^{-5}$). Importantly, in a control analysis substituting genetic loci with loci associated with type 2 diabetes[43], we did not observe a statistically significant overlap (ß =0.005, *P*=0.265) (see **Supplementary Table 14, Supplementary Figure 8**).

A look-up of left-handedness associated SNPs[13] in the methylation quantitative trait locus (mQTL) database by the Genetics of DNA Methylation Consortium (GoDMC)[44] showed that 95% of left-handedness associated SNPs associated with methylation levels of nearby (*cis;* 86 %) or distant (*trans;* 14%) CpGs (254 unique CpGs). We repeated the GWAS enrichment analysis with these CpGs driven by mQTLs removed, and obtained similar results, i.e. CpGs near GWAS loci (but not driven by significant mQTLs) were still more strongly associated with left-handedness compared to other genome-wide CpGs

(ß=0.027, *P*=0.027). None of the CpGs driven by mQTL was located in significant DMRs from our EWAS meta-analysis, or among the top 100 CpGs (by *p*-value) from our EWAS meta-analysis, which illustrates that our top EWAS findings are not driven by mQTL effects of the top GWAS loci.

Longitudinal analysis

While handedness is a stable trait, DNA methylation can vary over age[45]. We analyzed DNA methylation in ALSPAC offspring measured in cord blood at birth, and in peripheral blood at 7, 17, and 24 years old (*N*=791, **Table 2** and **Supplementary Table 3**) to examine the association between DNA methylation and left-handedness throughout childhood and adolescence. No associations survived adjustment for multiple tests at any time point (Bonferroni adjusted $P < 0.05$; **Supplementary Figure 9e-l**, **Supplementary Tables 19-26**). The correlations of effects for the top 100 CpG by p-value between time points were moderate to strong (mean correlation = 0.414; correlation range from *r*=0.355; *P*=0.0002 to *r*=0.578, $P=1.2\text{x}10^{-10}$), except for a weak correlation between top effects at 17 years and 24 years ($r^{ALSPAC17\text{-}ALSPAC24}$ =0.079; *P*=0.435) (**Supplementary Figure 10**). There were no overlapping CpGs amongst the top 100 CpGs between analyses at different time points (Supplementary Fig. 11). Correlations between top CpG effects between ALSPAC adults (mothers and fathers) and offspring at birth were strong negative ($r^{ALSPACadults\text{-}ALSPACatbirth}$ =-0.68; $P = 7.2\text{x}10^{-15}$) (**Supplementary Figure 12**), and between ALSPAC adults and offspring at 7, 17, 24 years were weak (r from -0.006 to 0.141, $P > 0.0003$) (**Supplementary Figure 10**).

DNA methylation in buccal cells

In NTR, buccal DNA methylation data (measured with the EPIC array at 787,711 CpG sites) were available in children (*N*=946, mean age 9.5, SD=1.85). The EWAS did not detect associations of DMPs with left-handedness (**Supplementary Figure 9m-n**, **Supplementary Table 27-28**). The effects for top 100 CpG in EWAS of handedness in buccal cells had weak correlations with effects in blood in the meta-analysis (from *r*=0.086, *P*=0.39 to *r*=0.179, *P*=0.07), NTR adults (from *r*=0.193, *P*=0.05 to *r*=0.268, *P*=0.007), ALSPAC adults (from *r*=-0.008, *P*=0.94 to *r*=-0.04, *P*=0.95), and ALSPAC offspring at different ages (from *r*=-0.384, $P=7.9\text{x}10^{-05}$ to *r*=0.312; *P*=0.002). Four DMRs in buccal cell DNA associated with left-handedness: DNA methylation was lower in left-handers at a DMR on chromosome 8 (4 CpGs; $P_{adj}= 9.14\text{x}10^{-06}$, average difference in DNA methylation in left-handers and right-handers 0.07%), a DMR on chromosome 9 (10 CpGs; P_{adj}=0.039, average difference 0.3%), a DMR on chromosome 12 (2 CpGs; P_{adj}=0.04, average difference 0.98%), and a DMR on chromosome 22 (2 CpGs; P_{adj}=0.035, average difference 1.1%) (**Table 4**, **Supplementary Figure 13**). These DMRs did not overlap with DMRs detected in the analyses of blood methylation data. Sixteen of 18 CpGs from these regions had a lower methylation level in left-handed children than in right-handed (**Supplementary Table 29**).

Sensitivity analyses

We reported the DNA methylation and left-handedness association study with adjustment for prenatal and postnatal factors that influence DNA methylation as shown in previous

TABLE 4. Significant differentially methylated regions associated with left-handedness in meta-analysis and secondary analysis

Cohort	Chromo-some	Start	End	*n* CpGs	Genes	Effect size	SE	*P*	P_{adjust}
Primary analysis									
Meta-analysis *NTR and ALSPAC adults (blood DNA methylation)*	chr20	36148679	36149022	16	*BLCAP NNAT*	-0.153	0.024	$9.80x10^{-11}$	$4.31x10^{-05}$
	chr2	9614471	9614744	7	*IAH1*	-0.102	0.019	$7.33x10^{-08}$	0.03
Secondary analysis									
NTR children at 9 years (buccal cell DNA methylation)	chr8	145024929	145025064	4	*PLEC1*	-0.056	0.008	$1.07x10^{-11}$	$9.14x10^{-06}$
	chr22	36011405	36011843	2	*MB*	-0.119	0.022	$4.09x10^{-08}$	0.035
	chr9	111696674	111697545	10	*EELP1, ABITRAM*	-0.134	0.024	$4.59x10^{-08}$	0.039
	chr12	899323	899559	2	*WNK1*	-0.117	0.021	$4.69x10^{-08}$	0.040

Effect size is a weighted sum of the EWAS effects for each CpG site in the DMR (i.e. methylation differences between LH and RH) where the weights account for dependence between CpG sites and uncertainty in the EWAS effects (see Methods). $N_{meta\text{-}analysis} = 3721$, $N_{NTR\ children} = 866$.
NTR, Netherlands Twin Register. ALSPAC, Avon Longitudinal Study of Parents and Children. SE, standard error; P_{adjust}, *P*-value multiplied by the total number of tests performed; the number of tests is equal to the number of regions for which DMR statistics are calculated.

studies: BMI[38] and smoking[46] in adults, and gestational age, birth weight and prenatal maternal smoking in children[47,48]. We examined if the EWAS results for handedness differ without taking these factors into account. Across all analyses, the correlations between the effects for the top 100 CpGs were strong between the models with and without adjustment for these factors (r ranged from 0.99 to 1), and overlaps of the top 100 CpGs were substantial (32 to 87 CpGs). Adjustment for the factors increased the number of DMRs associated with left-handedness in meta-analysis (1 DMR without adjustment and 2 DMRs with adjustment), and in EWAS in children (2 DMRs without adjustment, and 4 with adjustment in buccal cells in NTR).

Discordant MZ twins

In NTR, the DNA methylation datasets included 1,039 monozygotic (MZ) adult twins with blood samples (from the meta-analysis in NTR adults) and 794 MZ children with buccal samples (from NTR children in secondary analysis) from complete twin pairs with handedness data. We found that 21% of the MZ twin adult pairs (*N*=133 pairs) and 24% of the MZ twin child pairs (*N*=86 pairs) were discordant for handedness. Characteristics of MZ discordant twins are presented in Supplementary Table 4. In

both groups, we performed an MZ discordant within-pair EWAS analysis, comparing left- and right-handed twins. Within-pair analyses of DNA methylation of left- and right-handed twins did not identify DMPs or DMRs in blood or buccal samples at Bonferroni or FDR threshold (**Supplementary Figure 9a-d**, **Supplementary Tables 15-18**). We compared the methylation results obtained in discordant MZ twins to the EWAS meta-analysis results for the top 100 CpGs ranked on ascending p-value from each analysis. To avoid sample overlap, we repeated the EWAS meta-analysis after exclusion of the MZ discordant twin pairs. Correlations between the meta-analysis top 100 effects and mean methylation differences from the within-pair analysis were weak in adults ($r_{MZ\ disc\ adults\ blood–Meta\text{-}analysis}$=0.189, P=0.06; $r^{Meta\text{-}analysis–MZ\ disc\ adults\ blood}$=0.188, P=0.06, α = 0.0003) and children ($r_{MZ\ disc\ children\ buccal–Meta\text{-}analysis}$=0.134, P=0.18; $r^{Meta\text{-}analysis–MZ\ disc\ children\ buccal}$=0.252, P=0.01) (**Supplementary Figure 10**). There were few overlapping CpGs among the top 100 CpGs from the within-pair analyses and other analyses (**Supplementary Figure 11**).

Handedness methylation scores

To examine if variation in handedness can be predicted by DNA methylation levels across multiple CpGs, methylation scores (MS) were created. These were based on EWAS summary statistics in NTR to predict into ALSPAC, and on ALSPAC summary statistics to predict into NTR given the following p-value thresholds to include CpGs: $P<1x10^{-1}$, $P<1x10^{-3}$, $P<1x10^{-5}$. To estimate the variance explained by MS above genetic variants, polygenic scores (PGS) were created based on the summary statistics from the handedness GWAS of Cuellar-Partida et al.[13] with exclusion of NTR, ALSPAC and 23andMe cohorts (N_{GWAS}=196,419). Since four scores were tested (3 methylation scores, one polygenic score), we applied Bonferroni correction for four tests (α=0.05/4=0.0125) (**Supplementary Table 30**, **Supplementary Fig. 14**). The results are summarized in **Supplementary Table 31**. MS did not predict left-handedness in NTR and ALSPAC adults, or in children at 7-9 years old and did not explain variance over and above the variance explained by the PGS in the combined model (R^2_{MS} from -0.17 to 1.28%, R^2_{PGS} from 0.002 to 0.46%). The largest amount of explained variance was in ALSPAC at 7 years old for the MS based on CpGs at $P< 1x10^{-5}$ (R^2_{MS}=1.28%, P=0.1, N_{CpGs}=7).

Discussion

We have performed an epigenome-wide association study of left-handedness, including left- and right-handed individuals from two population-based cohorts from the Netherlands and the UK. In the meta-analysis, combining the NTR and ALSPAC adult cohorts, two DMRs associated with left-handedness. The first DMR (genomic location: chr20q11.23, 36,148,679:36,149,022) is located within the 5'UTR of the BLCAP apoptosis inducing factor *(BLCAP)* gene and nearby the transcription start site (TSS1500) of the neuronatin *(NNAT)* gene. BLCAP encodes a protein that reduces cell growth by stimulating apoptosis. *NNAT* is located within intron of BLCAP and is involved in brain development and nervous system structure maturation and maintenance. *BLCAP* and *NNAT* are imprinted in the brain[49]. The second intron of *NNAT* regulates the expression of *BLCAP* transcripts acting as an imprint control region regulating also allele-specific DNA methylation and histone

modifications[50]. Around 50% of CpGs in a region covering the *NNAT* CpG islands are methylated in brain and other tissues, suggestive of differential allele-specific methylation[51]. The imprinted DMR for these genes (chr20:36,139,941-36,159,190[52]) overlaps with the left-handedness DMR that we identified (chr20:36,148,679-36,149,022). A potential connection of genomic imprinting with handedness was previously suggested in a study of another imprinted gene *LRRTM1*[41]. CpGs from the handedness-associated region at chromosome 20 previously associated with myalgic encephalomyelitis and chronic fatigue syndrome, preterm birth, obesity, metabolic parameters, and arm fat mass (DXA scan measurement).

The second DMR (genomic location: chr2p25.1, 9,614,471:9,614,744) is located within the isoamyl acetate hydrolyzing esterase 1 *(IAH1)* gene. The *IAH1* gene encodes an acyl esterase and is associated with neonatal inflammatory skin and bowel disease, and a disease with an inborn error of leucine metabolism (3 methylglutaconic aciduria type 1). CpGs from the region previously associated with gestational age, bone mineral density, metabolic parameters, and schizophrenia. Some of these traits have been reported to be associated with handedness in epidemiological studies, e.g. BMI[53] and gestational age[54], for which we adjusted in our analyses. Previous analysis of the genetic correlations between left-handedness and 1,349 complex traits using LD-score regression did not reveal any genetic correlations at FDR <5%, but suggestive positive correlations were observed with neurological and psychiatric traits, including schizophrenia[13].

Even though no DMPs were identified after correction for multiple testing, and effect sizes of top CpGs were small (mean differences between left- and right-handed individuals smaller than 1%), the high-ranking CpGs are of potential interest. The second-ranking CpG cg09239756 (genomic location: chr12, 106,642,360) is located near the cytoskeleton associated protein 4 *(CKAP4)* gene. This gene mediates the anchoring of the endoplasmic reticulum to microtubules. Microtubules are an important cytoskeleton component that play a role in neuronal morphogenesis and migration, and axon transport[31]. Microtubules have been widely discussed in association with handedness[15,29] and brain anatomical asymmetry[32], and genes involved in microtubule pathways were enriched in the GWAS of handedness[13]. Moreover, in our enrichment analysis, we found that CpGs located within a 1Mb window from SNPs associated with left-handedness in the GWAS meta-analysis by Cuellar-Partida et al.[13] were more strongly associated with left-handedness in our meta-analysis compared to CpGs outside of this window. Larger EWAS meta-analysis or replication in additional independent cohorts is necessary to establish the robustness of the top DMPs.

Hand movements together with other lateralized movements and molecular signs of lateralization are observed at very early stages of human development in the uterus[6–8]. Therefore, DNA methylation differences associated with hand preference are expected to emerge early in development. While DNA methylation at some CpGs in the genome changes throughout the lifespan[45], the DNA methylation signal associated with left-handedness was moderately consistent from birth throughout the lifespan: DMP effects correlated in ALSPAC offspring from birth to 24 years old, although genome-wide significance for DMPs was not reached. Consistency in DNA methylation signal

associated with left-handedness at different time-points may indicate that the pattern for left-handedness is conserved through the lifespan.

Several DMRs were detected in buccal cells in children around 9 years old (genomic locations: chromosomes 8, 22, 9, and 12) after correction for multiple testing. Annotation of these regions implicate the following protein coding genes: the plectin gene *(PLEC1),* the myoglobin gene *(MB)* gene, the elongator complex protein 1 gene (ELP1), the actin binding transcription modulator gene *(ABITRAM),* and the WNK lysine deficient protein kinase 1 gene *(WNK1).* The CpGs in these regions are mostly hypomethylated in left-handed individuals. The genes encode for proteins that participate in cytoskeleton functions, chromatin organization, development of neurons, and metabolism. CpGs from DMRs in buccal cells previously associated with other phenotypes: CpGs from the DMR on chromosomes 8 with myalgic encephalomyelitis, chronic fatigue syndrome, multiple sclerosis, and gestational age; CpGs from the DMR on chromosome 9 with bone mineral density, tissue mass of the arm (DXA scans measurement), and multiple metabolic parameters. Interestingly, some of these phenotypes also associated with CpGs from our meta-analysis in blood DNA methylation.

A difference in handedness preference in MZ twin pairs has always fascinated parents of twins and twins themselves: how can children with almost identical genes differ for such a prominent trait? Handedness discordance in identical twins was described a long time ago[17,18], and the percentage of MZ discordant twins were reported as 20% of 3,486 MZ twins in East Flanders[19], and 19% of 1,724 MZ twins from a London twin study[28]. We observed that 21% of adult MZ twins and 24% of young MZ twins were discordant for handedness in our study, but we did not detect DNA methylation differences among them in blood or buccal cells. This null finding could mean that handedness discordance is not associated with methylation differences in the tissues that we studied (but might be present in other tissues), or that our analysis was underpowered to detect methylation differences associated with handedness discordance. Our discordant MZ twin analysis may be underpowered to detect small DNA methylation differences[55], as it included only 133 MZ discordant adult twin pairs and 86 child twin pairs. Different, not mutually exclusive, hypotheses have been proposed for handedness discordance of MZ twins, including large unique environmental effects, that are deduced based on the low heritability of around 25% as estimated in the meta-analysis by Medland et al.[10]. McManus[16] emphasizes that such unique environmental factors may represent what in biology is referred to as noise, randomness or fluctuating asymmetry. McManus quotes Mitchell[56] saying that such processes are “caused not by any factors outside the organism, but by inherent variation in the processes of development”. These reflect developmental variance unique to the person. In a discussion and review of the human and animal studies literature Molenaar et al.[57] called these processes ‘the third source of individual differences’, i.e. a third source besides genetic and environmental influences on individual differences and discuss how deterministic growth process may give rise to highly variable results.

There is a growing interest to improve the prediction of traits with use of other omics data than SNPs, like DNA methylation[37] and by non-genetic early life factors (e.g. earlier

factors associated with left-handedness including birth weight, being multiple, month of birth, breastfeeding etc.[58]). Given the low heritability of handedness (~25%[10,11]), it is expected that non-genetic factors play a role. Although together early life factors had minimal predictive value[58], they may inspire the search for DNA methylation signatures, as DNA methylation signatures later in life were found for birthweight[47]. Single CpGs did not individually reach statistical significance in our EWAS, but combining information across multiple CpGs into an overall methylation score can be a more powerful approach to capture variation in handedness. We calculated methylation scores as weighted sums of the individual's methylation loci beta values of a pre-selected number of CpG sites. However, the predictive value of polygenic and methylation scores for handedness was low, which likely reflects that curent GWAS and EWAS analyses for handedness are still underpowered.

Our multi-cohort epigenome-wide association study can be summarized in several key steps presented in **Figure 3**. We examined DNA methylation data in different tissues (whole blood, cord blood, buccal cells) and ages (from birth to adulthood). The limitations of the study are related to handedness measurements, available tissues, differences in platforms used for DNA methylation (Illumina 450k, EPIC), and study power. There were slight differences in the assessment of left-handedness between NTR and ALSPAC. Power might increase by investigating a refinement of the handedness phenotype (e.g. hand skill measurement rather than self-report of the preferred hand), analysis of DNA methylation in more relevant tissues, and an increase in sample size. The difference in left-handedness rates among children born before and after 1960 may be due to a move away from being forced to use the right hand prior to 1960[59]. We accounted for this trend by including age (which correlates almost perfectly with birth year in these samples) as a covariate in the analyses, however, it should be noted that the forced use of the right hand in older generations may render the phenotype definition of handedness less precise. Our meta-analysis was based on whole blood methylation data, where the methylation level represents the overall level of DNA obtained from millions of white blood cells. We observed methylation differences between left- and right-handed individuals of up to 0.8% at top-DMPs. This small effect size could reflect that a methylation difference is present in only a sub-set of cells in an individual, or a sub-set of individuals in the population, or a combination. The biological implications of these findings remain to be established and our top-DMPs remain to be replicated in additional cohorts or larger meta-analysis. The primary tissues of interest for handedness are brain[2,29], spinal cord[40], and arm muscle tissues4, and the timing when these tissues are collected could also play a role, but the collections of these tissues are not widely available in population cohorts for obvious reasons. Although 4,000 is considered a decent sample size for EWAS and smaller sample sizes have allowed for the successful detection of many loci where DNA methylation robustly associates with traits such as body mass index[38], the effect sizes for handedness were unknown a *priori* and 4,000 may still be too small to detect methylation differences associated with left-handedness. Similarly, sample sizes of EWASs of behavioral and psychiatric traits are now increasing beyond 10,000[60]. For GWAS of handedness that applied similar phenotyping as the current study, the increase of the sample size from about 400,000 to 1.7 million increased the number of associated loci from a handful to over forty[13], and similar increases in the number of detected loci may occur when EWAS sample sizes for left-handedness increase. Finally,

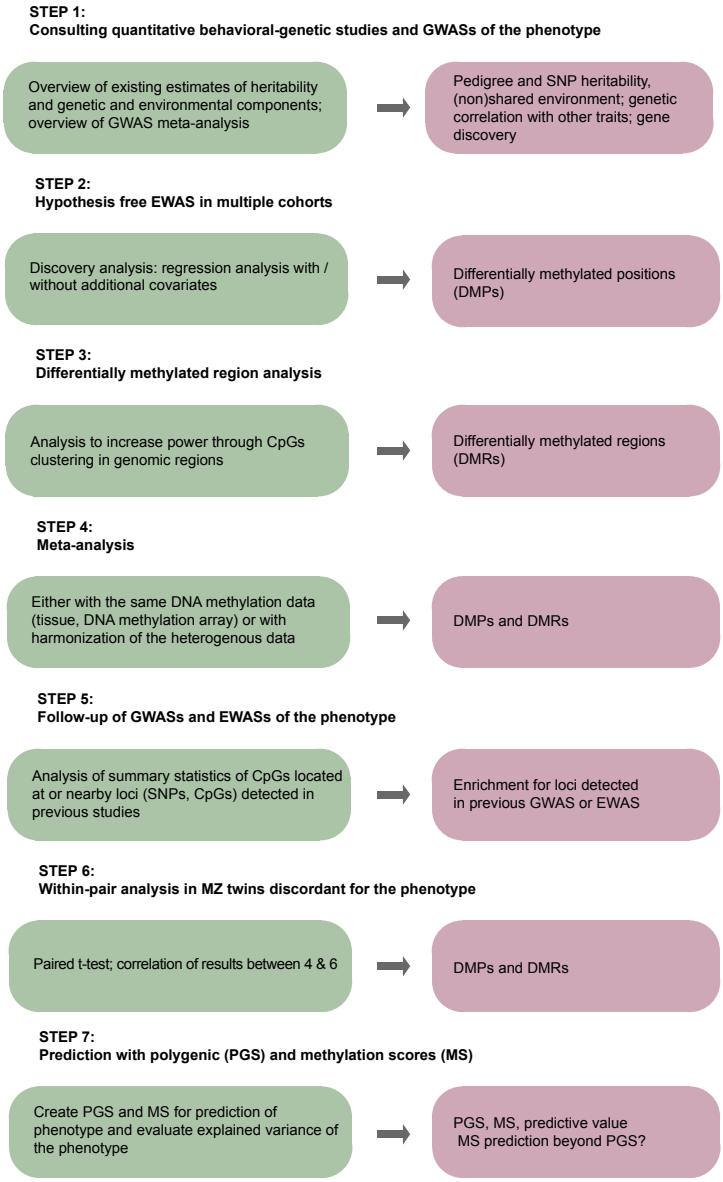


Figure 3. Seven-step approach to DNA methylation signature discovery, incorporating twin design

The figure represents the methodology of DNA methylation signature discovery study of a phenotype in multiple steps. It integrates behavioral-genetic and SNP-based methods (step 1) to estimate heritability, epigenome-wide study methods (steps 3-4) for association analyses, follow-up of results using summary statistics from previous EWASs and GWASs (step 5), the discordant twin design (step 6), and methods integrating polygenic and DNA methylation data (step 7) in enrichment analysis. Specific methods for each step are presented on the left and outcomes on the right. GWAS, genome-wide association study. EWAS, epigenome-wide association study. SNP, single nucleotide polymorphisms. CpG, cytosine-phosphate-guanine. PGS, polygenic scores. MS, methylation scores. DMPs, differentially methylated positions. DMRs, differentially methylated regions.

our study focused on DNA methylation, but other epigenetic processes could play a role in handedness such as histone modifications, post-translational regulation by miRNAs and X-chromosome inactivation that remain to be explored.

We reported an EWAS of left-handedness in two large population-based cohorts with data from children and adults, and examined performance of methylation scores and polygenic scores. Despite the plausible rationale of multiple genetic and non-genetic factors that may act via epigenetic pathways to influence the development of handedness, we did not uncover support for the hypothesis that DNA methylation in peripheral tissues captures much if any of the variation in handedness. We propose that future studies consider other tissues, such as related to central nervous system.

Methods

Overview. The primary epigenome-wide association study (EWAS) of left-handedness was performed in two cohorts with DNA methylation data in whole blood (Illumina, 450k): NTR adults[61] (*N*=2,682 individuals including twins, mean age at methylation 36.5, standard deviation (SD) 12.7), and ALSPAC adults[62,63] (*N*=1,232, mean age at methylation 48.98, SD 5.55). EWAS analyses were performed in each dataset separately, and summary statistics were combined in the meta-analysis (*N*=3,914) testing 409,563 CpGs. As this is a meta-analysis of existing DNA methylation datasets, no analyses were done to pre-determine sample sizes, but the sample size is larger compared to previously published DNA methylation studies of handedness[41,42]. We tested whether the EWAS signal was enriched in nearby loci detected in the previous GWAS on handedness[13]. Secondary analyses were performed in different tissues: in cord blood and peripheral blood in ALSPAC offspring[64], i.e. the children of ALSPAC participants that contributed to the primary EWAS (*N*=791 with DNA methylation data at birth, at 7, and 17 years old (Illumina 450k chip), and/or at 24 years old (Illumina EPIC array)), and in buccal cells from an independent group of children from the NTR[65,66] (*N*=946 twins, mean age 9.5, SD 1.85, Illumina EPIC array). The number with DNA methylation and covariate data in ALSPAC differed at different time points from 442 to 759. We carried out within-pair twin analysis in NTR MZ twins discordant for handedness (N_{adults}= 133 twin pairs, $N_{children}$= 86 twin pairs). We performed EWAS analyses in each dataset and examined correlations between the regression coefficients of top CpGs (ranked on ascending p-value) from each EWAS analysis. Finally, we created and tested polygenic and DNA methylation scores for left-handedness. The study method and design are presented in **Figure 1** and **Figure 3.** Detailed cohort information is provided in **Appendix 1**.

Handedness. *NTR.* Information on hand preference for adults and children was collected by surveys and in small subgroups from laboratory-based projects. Parental reports on children were collected at 5 years and included seven items for different activities, from which the item “What hand does child use for drawing?” was selected. The four answer categories were left-handed, right-handed, both hands and do not know. Multiple adult surveys included the question: “Are you right-handed or left-handed?” (4 surveys) or “Are you predominantly left-handed or right-handed?” (3 surveys). The three answer categories were left-handed, right-handed, and both. For a small number of adult self-

reports at younger ages (14, 16 or 18 years) or parental assessment at age 5 were also available.

ALSPAC. Adults (mothers and fathers) were asked which hand they used to write, draw, throw, hold a racket or bat, brush their teeth, cut with a knife, hammer a nail, strike a match, rub out a mark, deal from a pack of cards or thread a needle (11 questions). Child handedness was assessed at 42 months by questionnaire in which the mother was asked which hand the child used to draw, throw a ball, color, hold a toothbrush, cut with a knife, and hit things (6 questions). Responses were scored -1, 0 or 1 for left, either or right, respectively. Handedness was coded as 1 for left-handed or 0 for right-handed in both cohorts.

DNA methylation and genotyping. DNA methylation was measured with the Infinium HumanMethylation450 BeadChip Kit which measures more than 450,000 methylation sites (primary analysis in adults in NTR and ALSPAC and secondary analysis in ALSPAC offspring at birth, 7 and 17 years old), or the Infinium MethylationEPIC BeadChip which measures more than 850,000 methylation sites (secondary analysis in ALSPAC offspring at 24 years old and NTR children). Genotyping for polygenic risk scores was done on multiple platforms with imputation of the target data using reference haplotypes from 1000 genomes reference panel. Cohort-specific details on biosample collection, DNA methylation profiling, quality control, cell-type proportions measurements and genotyping are described in **Appendix 1** in **Supplementary material**.

Intergroup differences. We tested if there were differences in characteristics that were included in EWAS models (such as age at biological sample collection, sex, body mass index (BMI), smoking status at blood collection for adults, and gestational age, maternal smoking during pregnancy, birth weight for children, cell proportions/percentages in buccal swabs and in blood samples) between left- and right-handed individuals by generalized estimating equations (GEE) to accommodate the relatedness among the twins in NTR, and by standard logistic regression in ALSPAC. The R package 'gee' was used with the following specifications: binomial (for ordinal data) link function, 100 iterations, and the 'exchangeable' option to account for the correlation structure within families and within persons. Right- and lefthanded MZ discordant twins were compared with paired t-test for the traits that were not identical in twins (birth weight, BMI, smoking, cell percentages). All statistical tests here and below were two-tailed.

Epigenome-Wide Association Analyses. *Primary analyses.* The association between DNA methylation levels and left-handedness was tested for each site under a linear model (ALSPAC) or generalized estimating equation (GEE) model accounting for relatedness of twins (NTR). DNA methylation beta-values were the dependent variable and were typically normally distributed. The following predictors were included in the basic model: handedness (coded as 0=right-handed and 1=left-handed), sex, age at blood sampling, percentage of blood cells for blood samples, and technical covariates in NTR and ALSPAC (see **Appendix 1**). An adjusted model was fitted to account for BMI and smoking status at blood draw in both NTR and ALSPAC adult cohorts, because BMI and smoking have large effects on DNA methylation in adults[38,46]. The primary results reported in the paper are based on the fully adjusted model. The models are

described in **Appendix 2**. Throughout the text, we refer to regression coefficients from the EWAS, which represent the methylation difference between left-handed and right-handed individuals on the methylation beta-value scale. A positive regression coefficient (β) means a higher methylation level in left-handed individuals. The value of an individual on the methylation scale is commonly also symbolized as beta (β) and ranges from 0 to 1, where 0 represents a methylation level of 0% and 1 represents a methylation level of 100%.

Secondary analyses. The same basic models were fitted to the data from ALSPAC and NTR children. For DNA methylation in buccal cells, cell proportions (epithelial cells, natural killer cells) for buccal samples were included instead of percentage of blood cells. As several characteristics, such as gestational age and birthweight, affect DNA methylation[48,67], we included these in the adjusted model in children (see **Appendix 2** in **Supplementary material**).

In the within-pair analysis of discordant MZ twins, paired t-tests were employed to test for methylation differences between the left-handed and the right-handed twins. Paired t-tests were performed in R on residual methylation levels, which were obtained by adjusting the DNA methylation β-values for sample plate, array row, cell proportions in buccal samples in children and sample plate, array row, and percentages of blood cells in adults. Additional covariates, birth weight in children and BMI and smoking status in adults, were added in adjusted model. Age, sex, maternal smoking, and gestational age were not included because these variables are identical in MZ twins.

To account for multiple testing, we considered Bonferroni correction and a False Discovery Rate (FDR) of 5%. The Bonferroni corrected p-value threshold was calculated by dividing 0.05 by the number of genome-wide CpGs tested, and false discovery rate (FDR) q-values were computed with the R package 'qvalue' with default settings. The Bayesian inflation factor (λ) was calculated with the R package Bacon[64] (see **Supplementary Table 5**).

Meta-analysis. A meta-analysis was performed in METAL[69] based on estimates (regression coefficients) and standard errors from the EWAS of handedness performed with GEE in NTR and linear regression in ALSPAC. NTR and ALSPAC adult cohorts were combined. In total, 409,563 CpG sites present in both cohorts were tested with statistical significance evaluated after Bonferroni correction and at an FDR q-value <0.05.

Comparison of top CpGs from different analyses. To compare top CpGs from different analyses, we repeated the NTR EWAS analyses in adults in children and meta-analysis with discordant MZ twin pairs removed to avoid sample overlap. We selected methylation sites that overlapped in 13 analyses with adjusted model (meta-analysis, meta-analysis without discordant MZ twins, EWAS NTR adults, EWAS NTR adults without discordant MZ twins, EWAS ALSPAC adults, EWASs ALSPAC at birth, 7, 17, 24 years, EWAS NTR children, EWAS NTR children without discordant twins, and within-pair analyses of discordant MZ twin adults and children) that resulted in 379,924 methylation sites. We calculated Pearson correlations for effect estimates of the top 100 CpGs ranked by p-value from one analysis with the effect estimates of the same CpGs in other analyses. Statistical significance of correlations was assessed

after Bonferroni correction for the number of correlations tested: $\alpha = 0.05/(13 \times 13 - 13) = 0.0003$.

Differentially methylated regions. We used the R dmrff library[36] for R to identify regions where CpG sites showed evidence for association with handedness. Dmrff identifies DMRs by meta-analysing EWAS summary statistics from CpG sites in each region while adjusting for dependencies between the sites and uncertainty in the EWAS effects (https://github.com/perishky/dmrff). In this study, dmrff was applied separately in the NTR and ALSPAC cohorts, and then used to identify DMRs in common between the cohorts by meta-analysis. As previously described[36], DMR meta-analysis preceded by first identifying candidate DMRs using the EWAS meta-analysis summary statistics, calculating DMR statistics for these candidates in each cohort separately, and then meta-analysing the DMR statistics across the two cohorts. The DMR effect size is a weighted sum of the EWAS effects for each CpG site (i.e. methylation differences between LH and RH). All dmrff p-values were adjusted for multiple tests (Bonferroni adjustment) by multiplying them by the total number of DMRs considered. We report significant regions ($P_{adj} < 0.05$) including at least two CpG sites within a 500bp window observed to be nominally associated with handedness by EWAS ($P < 0.05$). The average absolute DNA methylation difference in the region between left-handers and right-handers is calculated as the sum of absolute regression coefficients of each CpG in the region divided by the number of CpGs. We plotted the DMRs with the coMET R Bioconductor package[70] to graphically display additional information on physical location of CpGs, correlation between sites, statistical significance, and functional annotation (annotation tracks included genes Ensembl, CpG islands (UCSC), regulation Ensembl).

GWAS follow-up. GWAS follow-up analyses were performed to examine whether CpGs within a 1 Mb window of loci detected by the GWAS for left-handedness[13], on average, showed a stronger association with left-handedness than other genome-wide methylation sites (Infinium HumanMethylation450 BeadChip). We obtained a SNP list based on the GWAS meta-analysis without NTR, ALSPAC, and 23andMe by Cuellar-Partida et al.[13] (196,419 individuals, $N_{SNPs} = 13{,}550{,}404$), from which we selected all SNPs with a p-value $<1.0 \times 10^{-08}$, $<1.0 \times 10^{-06}$, and $<1.0 \times 10^{-05}$, and determined the distance of each Illumina 450k methylation site to each SNP. To test whether methylation sites near GWAS loci were more strongly associated with left-handedness, meta-analysis EWAS test statistics were regressed on a variable indicating if the CpG is located within a 1 Mb window from SNPs associated with handedness (1=yes, 0=no):

$$|Zscore| = Intercept + \beta_{category\ x} * Category\ x,$$

where |Zscore| represents the absolute Zscore for a CpG from the EWAS meta-analysis of handedness; $\beta_{category\ x}$ represents the estimate for category x, i.e. the change in the EWAS test statistic associated with a one-unit change in category x (e.g. being within 1 Mb of SNPs associated with left-handedness). For each enrichment test, bootstrap standard errors were computed with 2000 bootstraps with the R-package “simpleboot”. Statistical significance was assessed at $\alpha=0.05$. As control analysis, the same follow-up was performed using GWAS summary statistics on a trait that is unrelated to handedness – type 2 diabetes in UK Biobank cohort (N=244,890)[43]. GWAS summary statistics were

downloaded from GWASAtlas (https://atlas.ctglab.nl/traitDB/3686; 41204_E11_logistic.EUR.sumstats.MACfilt.txt; accessed on February 1 2021).

We looked up CpG sites associated with handedness-associated SNPs with a p-value $<1.0\times10^{-08}$ (based on the GWAS meta-analysis by Cuellar-Partida et al.[13] without 23andMe, NTR and ALSPAC, resulting in a list of 420 SNPs) using the mQTL database maintained by the Genetics of DNA Methylation Consortium (GoDMC, *N*=27,750 European samples; http://mqtldb.godmc.org.uk/about)[44]. We then checked if associations with handedness were observed for these sites in our EWAS meta-analysis and DMR meta-analysis.

EWAS follow-up. To examine previously reported associations for epigenome-wide significant DMRs associated with left-handedness in our study, we looked up CpGs from the regions in the EWAS Atlas[71] (https://bigd.big.ac.cn/ewas/tools; accessed on August 1 2020) and EWAS catalogue[72] (http://www.ewascatalog.org; access on November 1 2020).

Polygenic and methylation scores. *Polygenic scores* (PGS) for handedness were calculated based on the GWAS meta-analysis without 23andMe by Cuellar-Partida et al.[13]. To avoid overlap between the discovery and target samples, summary statistics without NTR and ALSPAC were requested (196,419 individuals, N_{SNPs} = 13,550,404). The linkage disequilibrium (LD) weighted betas were calculated using a LD pruning window of 250 KB, with the fraction of causal SNPs set at 0.50 by LDpred[73]. We randomly selected 2500 2nd degree unrelated individuals from each cohort as a reference population to calculate the LD patterns. The resulting betas were used to calculate the PGSs in each dataset using the PLINK 1.9 software. All PGSs were standardized (mean of 0 and standard deviation of 1). *Methylation scores* (MS) were calculated in NTR based on EWAS summary statistics obtained from ALSPAC, and vice versa, as previously done to create methylation scores for BMI and height[74]. We calculated same-tissue same-age DNA-methylation scores based on methylation data from NTR adults (blood) and ALSPAC parents (blood), and cross-tissue DNA-methylation scores based on data from NTR and ALSPAC offspring, with DNA methylation measured in buccal cells, and blood, respectively (see **Figure 1**). For each individual, a weighted score sum was calculated for left-handedness by multiplying the methylation value at a given CpG by the effect size of the CpG (β), and then summing these values over all CpGs:

$$\textit{DNA methylation score (i)} = \beta_1 * CpG_{1i} + \beta_2 * CpG_{2i} \ldots + \beta_n * CpG_{ni},$$

where CpGn is the methylation level at CpG site *n* in participant *i*, and β_n is the regression coefficient at CpGn taken from summary statistics of the EWAS analysis. All methylation scores were standardized (mean of 0 and standard deviation of 1). We used weights from summary statistics of EWASs in four cohorts: NTR adults, ALSPAC adults, NTR children, ALSPAC offspring at 7 years old. Subsets of CpGs to be included in methylation scores were selected based on p-value $<1x10^{-1}$, $<1x10^{-3}$, and $<1x10^{-5}$. We analysed the predictive value of the left-handedness polygenic scores and methylation scores in NTR and ALSPAC adult and child cohorts from our EWAS study. To quantify the variance explained by the PGS and MS, we used the approach proposed by Lee et al.[75], where coefficients of determination (R^2) for binary responses are calculated on the liability scale. The equations of all models are provided in **Appendix 2** in **Supplementary material**.

Statistical significance was assessed following Bonferroni correction for the number of scores tested (PGS and 3 MSs). This resulted in $\alpha=0.05/4=0.0125$, nominal significance at 0.05.

Ethics statement. All methods were performed in accordance with the Declaration of Helsinki. For NTR, the study was approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Centre, Amsterdam, an Institutional Review Board certified by the U.S. Office of Human Research Protections (IRB number IRB00002991 under Federal-wide Assurance FWA00017598; IRB/institute codes, NTR 03-180). All subjects provided written informed consent. For children, written informed consent was given by their parents. For *ALSPAC*, ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. All subjects provided written informed consent. For children, written informed consent was given by their mothers.

References

1. Willems, R. M. *et al.* On the other hand: including left-handers in cognitive neuroscience and neurogenetics. Nature reviews. *Neuroscience* vol. 15 193–201 (Nature Publishing Group, 2014).
2. De Gennaro, L. *et al.* Handedness is mainly associated with an asymmetry of corticospinal excitability and not of transcallosal inhibition. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology* **115**, 1305–1312 (2004).
3. Sarkar, A., Dutta, S., Bal, K. & Biswas, J. Handedness may be related to variations in palmar arterial arches in humans. *Singapore medical journal* **53**, 409–412 (2012).
4. Diederichsen, L. P. *et al.* The effect of handedness on electromyographic activity of human shoulder muscles during movement. *Journal of electromyography and kinesiology: official journal of the International Society of Electrophysiological Kinesiology* **17**, 410–419 (2007).
5. Stoyanov, Z., Decheva, L., Pashalieva, I. & Nikolova, P. Brain asymmetry, immunity, handedness. Cent.Eur.J.Med 7 (1), 1–8 (2012).
6. Hepper, P. G. The developmental origins of laterality: Fetal handedness. *Developmental Psychobiology* **55**, 588–595 (2013).
7. De Vries, J. I. P. *et al.* Fetal handedness and head position preference: A developmental study. *Developmental Psychobiology* **39**, 171–178 (2001).
8. Parma, V., Brasselet, R., Zoia, S., Bulgheroni, M. & Castiello, U. The origin of human handedness and its role in pre-birth motor control. Scientific Reports 7, 16804 (2017).
9. Papadatou-Pastou, M. *et al.* Human handedness: A meta-analysis. *Psychological bulletin* **146**, 481–524 (2020).
10. Medland, S. E., Duffy, D. L., Wright, M. J., Geffen, G. M. & Martin, N. G. Handedness in Twins: Joint Analysis of Data From 35 Samples. *Twin Research and Human Genetics* **9**, 46–53 (2006).
11. Medland, S. E. *et al.* Genetic influences on handedness: data from 25,732 Australian and Dutch twin families. *Neuropsychologia* **47**, 330–337 (2009).
12. Vuoksimaa, E., Koskenvuo, M., Rose, R. J. & Kaprio, J. Origins of handedness: a nationwide study of 30,161 adults. *Neuropsychologia* **47**, 1294–1301 (2009).
13. Cuellar-Partida, G. *et al.* Genome-wide association study identifies 48 common genetic variants associated with handedness. *Nature human behaviour* **5**, 59–70 (2021).
14. Annett, M. Left, right, hand, and brain: the right shift theory. (L. Erlbaum Associates, 1985).
15. McManus, C. Half a century of handedness research: Myths, truths; fictions, facts; backwards, but mostly forwards. *Brain and Neuroscience Advances* **3**, 239821281882051 (2019).
16. McManus, C. Is any but a tiny fraction of handedness variance likely to be due to the external environment? Laterality 1–5 (2021).
17. Rife, D. C. Handedness, with special reference to twins. *Genetics* **178–186** (1940).
18. Segal, N. L. Twin Mythconceptions: False Beliefs, Fables, and Facts about Twins. (2017).
19. Derom, C., Thiery, E., Vlietinck, R., Loos, R. & Derom, R. Handedness in Twins According to Zygosity and Chorion Type: A Preliminary Report. Behavior Genetics vol. 26(4), 407-408 (1996).
20. Francks, C. *et al.* LRRTM1 on chromosome 2p12 is a maternally suppressed gene that is associated paternally with handedness and schizophrenia. *Molecular psychiatry* **12**, 1057,1129-1139 (2007).
21. Sun, T. *et al.* Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science (New York, N.Y.)* **308**, 1794–1798 (2005).
22. Arning, L. *et al.* PCSK6 VNTR Polymorphism Is Associated with Degree of Handedness but Not Direction of Handedness. *PloS one* **8**, e67251 (2013).
23. Brandler, W. M. *et al.* Common variants in left/right asymmetry genes and pathways are associated with relative hand skill. P*LoS genetics* **9**, e1003751 (2013).

24. Arning, L. *et al.* Handedness and the X chromosome: the role of androgen receptor CAG-repeat length. *Scientific reports* **5**, 8325 (2015).
25. Hampson, E. & Sankar, J. S. Hand preference in humans is associated with testosterone levels and androgen receptor gene polymorphism. Neuropsychologia 50(8), 2018–2025 (2012).
26. Medland, S. E. *et al.* Opposite Effects of Androgen Receptor CAG Repeat Length on Increased Risk of Left-Handedness in Males and Females. *Behavior Genetics* **35**, 735–744 (2005).
27. Eriksson, N. *et al.* Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS genetics* **6**, e1000993 (2010).
28. Armour, J. al, Davison, A. & McManus, I. C. Genome-wide association study of handedness excludes simple genetic models. *Heredity* **112**, 221–225 (2014).
29. Wiberg, A. *et al.* Handedness, language areas and neuropsychiatric diseases: insights from brain imaging and genetics. *Brain: a journal of neurology* **142**, 2938–2947 (2019).
30. De Kovel, C. G. F. & Francks, C. The molecular genetics of hand preference revisited. *Scientific Reports* **9**, 5986 (2019).
31. Conde, C. & Cáceres, A. Microtubule assembly, organization and dynamics in axons and dendrites. *Nature Reviews Neuroscience* **10**, 319–332 (2009).
32. Sha, Z. *et al.* The genetic architecture of structural left-right asymmetry of the human brain. *Nature human behaviour* **5(9)**, 1226–1239 (2021).
33. Klar, A. J. S. An Epigenetic Hypothesis for Human Brain Laterality, Handedness, and Psychosis Development. *Cold Spring Harbor Symposia on Quantitative Biology* **69**, 499–506 (2004).
34. Crow, T. J. A theory of the origin of cerebral asymmetry: epigenetic variation superimposed on a fixed right-shift. *Laterality* **15**, 289–303 (2010).
35. Schmitz, J., Metz, G. A. S., Güntürkün, O. & Ocklenburg, S. Beyond the genome-Towards an epigenetic understanding of handedness ontogenesis. *Progress in Neurobiology* **159**, 69–89 (2017).
36. Suderman, M. *et al.* dmrff: identifying differentially methylated regions efficiently with power and control. bioRxiv 508556 (2018) doi:10.1101/508556.
37. Hüls, A. & Czamara, D. Methodological challenges in constructing DNA methylation risk scores. *Epigenetics* **15**, 1–11 (2020).
38. Wahl, S. *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* **541**, 81–86 (2017).
39. Schmitz, J., Güntürkün, O. & Ocklenburg, S. Building an Asymmetrical Brain: The Molecular Perspective. *Frontiers in Psychology* **10**, 982 (2019).
40. Ocklenburg, S. *et al.* Epigenetic regulation of lateralized fetal spinal gene expression underlies hemispheric asymmetries. eLife 6, e22784 (2017).
41. Leach, E. L., Prefontaine, G., Hurd, P. L. & Crespi, B. J. The imprinted gene LRRTM1 mediates schizotypy and handedness in a nonclinical population. *Journal of Human Genetics* **59**, 332–336 (2014).
42. Schmitz, J., Kumsta, R., Moser, D., Güntürkün, O. & Ocklenburg, S. DNA methylation in candidate genes for handedness predicts handedness direction. *Laterality* **23**, 441–461 (2018).
43. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nature genetics* 51, 1339–1348 (2019).
44. Min JL *et al.* Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat Genet.* 53(9):1311-1321 (2021).
45. Mulder, R. H. *et al.* Epigenome-wide change and variation in DNA methylation in childhood: Trajectories from birth to late adolescence. *Human molecular genetics* (2021) doi:10.1093/hmg/ddaa280.

46. Joehanes, R. *et al.* Epigenetic Signatures of Cigarette Smoking. Circulation. *Cardiovascular genetics* **9**, 436–447 (2016).
47. Küpers, L. K. *et al.* Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. Nature communications, 10(1), 1893 (2019).
48. Joubert, B. R. *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *The American Journal of Human Genetics* **98**, 680–696 (2016).
49. Schulz, R. *et al.* Transcript- and tissue-specific imprinting of a tumour suppressor gene. *Human molecular genetics* **18**, 118–127 (2009).
50. Thamban, T. *et al.* The putative Neuronatin imprint control region is an enhancer that also regulates the Blcap gene. *Epigenomics* **11**, 251–266 (2019).
51. Evans, H. K., Weidman, J. R., Cowley, D. O. & Jirtle, R. L. Comparative phylogenetic analysis of blcap/nnat reveals eutherian-specific imprinted gene. *Molecular biology and evolution* **22**, 1740–1748 (2005).
52. Court, F. *et al.* Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome research* **24**, 554–569 (2014).
53. Propper, R., Struble, C. & Brunyé, T. Handedness and Physical Measures II: Objectively Measured Height and Weight. *International Journal of School and Cognitive Psychology* **2**, 1000127 (2015).
54. Domellöf, E., Johansson, A.-M. & Rönnqvist, L. Handedness in preterm born children: a systematic review and a meta-analysis. *Neuropsychologia* **49**, 2299–2310 (2011).
55. Tsai, P.-C. & Bell, J. T. Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. *International journal of epidemiology* **44**, 1429–1441 (2015).
56. Mitchell, K. J. *Innate.* (Princeton University Press, 2018).
57. Molenaar, P. C., Boomsma, D. I. & Dolan, C.V. A third source of developmental differences. *Behavior genetics* **23**, 519–524 (1993).
58. De Kovel, C. G. F., Carrión-Castillo, A. & Francks, C. A large-scale population study of early life factors influencing left-handedness. *Scientific reports* **9**, 584 (2019).
59. Porac, C. & Friesen, I. C. Hand preference side and its relation to hand preference switch history among old and oldest-old adults. *Developmental neuropsychology* **17**, 225–239 (2000).
60. Van Dongen, J. *et al.* DNA methylation signatures of aggression and closely related constructs: A meta-analysis of epigenome-wide studies across the lifespan. *Molecular psychiatry* **26(6)**, 2148–2162 (2021)
61. Willemsen, G. *et al.* The Netherlands Twin Register Biobank: A Resource for Genetic Epidemiological Studies. *Twin Research and Human Genetics* **13**, 231–245 (2010).
62. Boyd, A. *et al.* Cohort Profile: The 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology* **42**, 111–127 (2013).
63. Fraser, A. *et al.* Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology* **42**, 97–110 (2013).
64. Northstone, K. *et al.* The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. *Wellcome open research* **4**, 51 (2019).
65. Boomsma, D. I. Aggression in children: unravelling the interplay of genes and environment through (epi)genetics and metabolomics. Journal of Pediatric and Neonatal Individualized Medicine (JPNIM) 4, e040251–e040251 (2015).
66. Bartels, M. *et al.* Childhood aggression and the co-occurrence of behavioural and emotional problems: results across ages 3–16 years from multiple raters in six cohorts in the EU-ACTION project. *European Child & Adolescent Psychiatry* **27**, 1105–1121 (2018).

67. Küpers, L. K. *et al.* DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *International journal of epidemiology* **44**, 1224–1237 (2015).

68. Van Iterson, M., van Zwet, E. W. & Heijmans, B. T. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biology* **18**, 19 (2017).

69. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics (Oxford, England)* **26**, 2190–1 (2010).

70. Martin, T. C., Yet, I., Tsai, P.-C. & Bell, J. T. coMET: visualisation of regional epigenome-wide association scan results and DNA co-methylation patterns. *BMC Bioinformatics* **16**, 131 (2015).

71. Li, M. *et al.* EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Research* **47**, D983–D988 (2019).

72. Battram, T. *et al.* The EWAS Catalog: a database of epigenome-wide association studies (accessed August, 2021) doi:10.31219/OSF.IO/837WN.

73. Vilhjálmsson, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. A*merican journal of human genetics* **97**, 576–92 (2015).

74. Shah, S. *et al.* Improving Phenotypic Prediction by Combining Genetic and Epigenetic Associations. *The American Journal of Human Genetics* **97**, 75–85 (2015).

75. Lee, S. H., Goddard, M. E., Wray, N. R. & Visscher, P. M. A better coefficient of determination for genetic profile analysis. *Genetic epidemiology* **36**, 214–224 (2012).

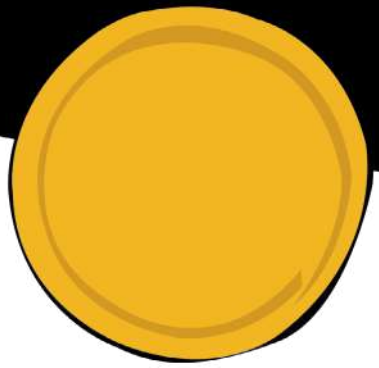
Supplements

The supplementary materials are available online at

https://doi.org/10.1038/s41598-022-08998-0

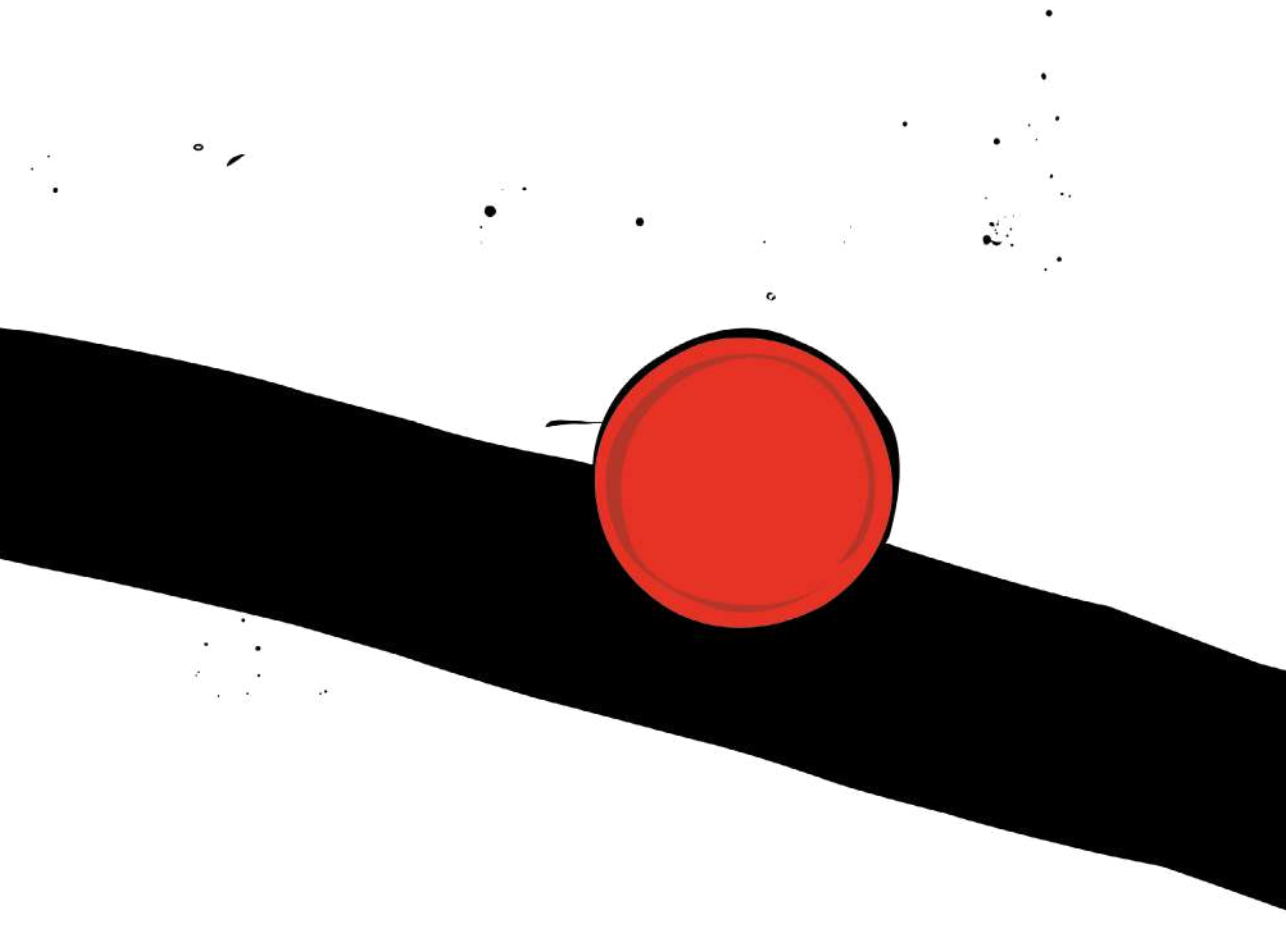
PART II

Genetic and epigenetic prediction of complex traits



Chapter 7

Genomics of human aggression: current state of genome-wide studies and an automated systematic review tool



Abstract

There are substantial differences, or variation, between humans in aggression, with its molecular genetic basis mostly unknown. This review summarizes knowledge on the genetic contribution to variation in aggression with three foci: (1) a comprehensive overview of reviews on the genetics of human aggression, (2) a systematic review of genome-wide association studies (GWASs), and (3) an automated tool for the selection of literature based on supervised machine learning.

The phenotype definition "aggression" (or "aggressive behaviour", or "aggression-related traits") included anger, antisocial behaviour, conduct disorder, and oppositional defiant disorder. The literature search was performed in multiple databases, manually and using a novel automated selection tool, resulting in 18 reviews and 17 GWASs of aggression.

Heritability estimates of aggression in children and adults are around 50%, with relatively small fluctuations around this estimate. In 17 GWASs, 817 variants were reported as suggestive ($P \leq 1.0 \times 10^{-05}$), including 10 significant associations ($P \leq 5.0 \times 10^{-08}$). Nominal associations ($P \leq 1 \times 10^{-05}$) were found in gene-based tests for genes involved in immune, endocrine, and nervous systems. Associations were not replicated across GWASs. A complete list of variants and their position in genes and chromosomes are available online. The automated literature search tool produced literature not found by regular search strategies. Aggression in humans is heritable, but its genetic basis remains to be uncovered. No sufficiently large genome-wide association studies have been carried out yet. With increases in sample size, we expect aggression to behave like other complex human traits for which GWAS has been successful.

In 2021 the study was updated with two reviews, a new study of GWAS, review of epigenome-wide association studies and omics domain prepared for publication Odintsova VV, Hagenbeek FA, van der Laan CM, van de Weijer S., Boomsma DI. Genetics and epigenetics of human aggression. In: Handbook of Clinical Neurology. Brain and Crime (Eds. Swaab J.T.). Elsevier (2022) (in press). The updates are indicated in the text.

Published as: Odintsova V.V.*, Roetman P.J*, Pool R., Ip H. F., Van der Laan C.M., Tona D.K., Vermeiren R.R.J.M., Boomsma D. I. (2019). Genomics of human aggression: current state of genome-wide studies and an automated systematic review tool. *Psychiatric Genetics,* 29(5): 170-190.

Introduction

Aggression is a common type of human behaviour[1] and is considered a characteristic that is shared by all humans[2]. The propensity for aggression, however, varies considerably between individuals. This paper addresses the question to what extent the variation that is seen for aggression has a genetic cause. Broadly, aggression can be defined as a behaviour that intends to cause physical or emotional harm to others[3]. High levels of aggression are also seen in individuals with severe mental disorders (e.g., autism, bipolar disorder, schizophrenia) as well as in patients with (rare) Mendelian disorders[4]. Because of the large impact of aggression on the affected individual, their families, their environment, and society as a whole, there is a substantial interest in studying aggression from a wide range of disciplines. In this context, one goal is to unravel the aetiology of aggression by identifying environmental exposures and biomarkers, including genetic factors, epigenetic marks, and metabolites, that could function as predictors of (excessive) aggression[5].

Research often focuses on the pathological aspects of aggressive behaviour, while aggression does not solely have negative consequences or outcomes. Under certain circumstances, aggressive behaviour is beneficial to individuals, for example when competing for limited resources, like food or mates[6], or achieving social dominance[7]. Aggression can further be a powerful deterrent against aggressive behaviour from others. Because both high and low levels of aggression can be detrimental to survival and procreation, it has been postulated that aggression is under stabilizing selection, implying that variation in aggression should show significant heritability. Substantial heritability estimates have indeed been reported in animals[8] and humans, as reviewed below.

Benefits of aggressive acts depend on the type of aggression, its success, environmental circumstances and also vary across cultures[9]. For example, predatory goal-oriented aggression has been associated with social dominance in some instances[10–12], but this association seems to vary between groups that are more prosocial and groups that consist predominantly of individuals with disruptive behaviour problems[13]. A decrease in social status can also result from aggression, in particular from reactive aggression, which is an uncontrolled type of aggression stemming from internal or external frustration. In reverse, after a conflict, proactive aggression is increased in the victorious party while the losing party is less likely to engage in another aggressive act[14,15]. To differentiate between different outcomes of aggression, researchers have distinguished aggression subtypes (e.g. reactive vs. proactive; overt vs. covert), developmental stages (childhood vs. adolescent onset), and comorbidities (e.g., with internalizing problems or with attention deficit hyperactivity disorder (ADHD)). In summary, the outcomes and types of aggressive acts can differ greatly between persons and circumstances, and need not always be dysfunctional.

At the start of the 1990s, research on aggressive behaviour was given a new impulse by a seminal paper of Brunner et al.[16], in which a Dutch pedigree was described where men exhibited impulsive aggression, arson, violence, and borderline mental retardation. The family appeared to have a rare point mutation in the structural gene for monoamine-oxidase-A *(MAOA)* – which codes for an enzyme that is involved in the oxidative deamination of neurotransmitters like dopamine, serotonin and norepinephrine – resulting in a deficiency of the MAOA enzyme. A study, by Caspi et al.[17], compared variants of the *MAOA* gene in children who experienced maltreatment and showed that children with

the variant resulting in lower levels of the MAOA enzyme were more likely to develop antisocial behavior (ASB). Efforts to replicate the latter finding have been contradictory, either without replication[18,19] or with replication[20–22]. Nevertheless, the studies of Brunner and Caspi stressed the importance of biological factors in the development of aggression and ASB. This instigated extensive efforts to study the genetic basis of aggression.

Enormous progress has been made with respect to technology in molecular biology and large-scale genotyping, as well as in the development of statistical methods for genetic association studies and polygenic scores for individual risk assessment, once sufficiently large genetic-association studies are available[23]. Costs for genotyping and sequencing of DNA, the epigenome and of RNA, and biomarker assessment, such as metabolomics, have steadily decreased, allowing for large studies, relating aggressive behaviour to genome, epigenome, transcriptome, and other biomarkers[24]. Progress also has been made in characterizing the exposome, which reflects the totality of a person's environmental exposures in space and time[25].

Genome-wide association studies (GWAS) provide a conceptual framework to examine whether individual differences in aggression are associated with allelic differences in millions of single nucleotide polymorphisms (SNPs) across the genome[26]. Because a GWAS targets the entire human genome, it enables a data-driven approach to identify loci of interest. This hypothesis-free approach could potentially help researchers to overcome limits imposed by multifactorial nature of a trait and incomplete understanding of its physiological basis.

Here we synthesise knowledge deriving from studies on genetics of human aggression and variance in liability to aggression-related traits. Our review has three foci: (1) to give a comprehensive overview of reviews already done on genetics of human aggression, (2) to carry out a systematic review of GWAS studies on human aggression, and (3) to introduce an automated systematic review for the selection of relevant literature, based on supervised machine learning. For consistency, in this review, we will use the general term 'aggression' (or 'aggressive behaviour', or 'aggression-related traits') to refer to the terminologies used by different authors (see **Box 1**), including anger, hostility dimensions, parent-reported child aggressive behaviour, physical aggression, antisocial behaviour (ASB), violent offending, conduct disorders (CD), oppositional defiant disorder (ODD), and antisocial personality disorder (ASPD).

Methods

To optimize detection of the relevant literature for our review, we incorporated two strategies:

1. A 'traditional' (manual) search strategy where search terms were used to extract the relevant articles from literature databases.

2. An automated screening with Automated Systematic Review Software (ASR) where relevant articles were detected via the utilization of machine learning algorithms and a software development platform.

Box 1. Definitions of Aggression

Concept	Definitions
Reactive/hostile/ affective/impulsive aggression Proactive/ instrumental/pre-mediated aggression	Angry or frustrated responses to a real or perceived threat (Tuvbald and Baker, 2011) Aggressive response to a perceived threat or provocation (Waltes, 2016) Planning, the motive of the act extends beyond harming the victim (Tuvbald and Baker, 2011) Planned antisocial behavior that anticipates a reward or dominance over others (Waltes, 2015)
Indirect/relational aggression	Intentionally causing pain or harm to the victim (Tuvblad and Baker, 2011) Relational social manipulation such as gossip and peer exclusion (Tuvblad and Baker, 2011)
Chronic physical aggression	Tendency to use physical aggression more frequently than the large majority of a birth cohort over many years (Tremblay et al., 2018; Provencal et al., 2015)
Externalizing behavior	Behavior that directs problematic energy outward and is expressed as aggression, defiance, bullying, vandalism, theft, and other socially unacceptable actions (Anholt and Mackay, 2012)
Aggression and anger-related traits associated with suicidal behavior	Anger can be conceptualized as a core construct of related traits or variables inwardly and/or outwardly expressed such as aggression, rage, and hostility (Spielberger et al, 1985 cite: Baud, 2005) Aggression and anger-related traits are considered risk factors for suicidal behavior (Baud, 2005)
Aggression related phenotype	A dimensional trait including externalizing behavior, anger, delinquency, criminality, violence or a diagnostic category (conduct disorder, oppositional defiant disorder, callous unemotional, and antisocial personality) (Fernandez-Castillo and Cormand, 2016)
Frustrative non-reward aggression Defensive aggression Offensive (or proactive) aggression	Behaviors that correspond to the withdrawal or prevention of reward Behaviors caused by the perception of an immediate threat, which have the goal of eliminating the threat Instrumental behaviors aimed at achieving a positive goal, often in the face of competition or in the context of social hierarchies (RDoC (National Institute of Mental Health); Veroude et al., 2016)
Aggression as behavior category in conduct disorder (CD)	CD is a developmental disorder characterized by a consistent pattern of externalizing behavior, developing during childhood or adolescence, where an individual displays aggression toward people or animals, destroys property, exhibits deceit by lying or stealing, and/or seriously violates societal rules or norms (DSM-V) Conduct disorder is a psychiatric disorder of childhood and adolescence characterized by aggression toward people and animals, destruction of property, deceitfulness or theft, and serious violation of rules (DSM-V (American Psychiatric Association, 2013); Salvatore et al., 2018)
Antisocial behavior	Refers to actions that violate social norms in ways that reflect the violation of others' rights (Moffit, 2005; Gard et al., 2018)
Aggression as violence	*No definition is given* (Vassos et al., 2014)

Traditional approach

Search strategy

Search terms were developed by the authors based on prior literature and discussions with an expert librarian (J.W.S) from the LUMC. A literature search was performed in PubMed, Embase, Web of Science, Cochrane library, PsychInfo and Academic Search Premier with a comprehensive list of general search terms and medical subject headings (**Supplement S2**). Searches were conducted separately for reviews/meta-analyses and GWA studies. Searches included literature without a specific time limit and were conducted in mid-April 2019.

Selection criteria

A selection was made from all titles and abstracts that were found in the databases using pre-specified inclusion and exclusion criteria (**Table 1**). Articles were included if they (1) were written in English and (2) focused on human aggression. Studies were excluded if (1) they focused on animals, or (2) general terms linked to 'aggression/violent etc.' did not refer to a psychological/ psychiatric perspective but rather to characteristics of disease (e.g. aggressive cancer), or (3) articles discussed only a single gene. Psychiatric disorders, which incorporate acts of aggression and are highly correlated to aggression and antisocial lifestyles, like ODD, CD, and ASPD, were included. Articles referring to associations between genetic data and other (neuro)psychiatric disorders as main outcome (e.g. psychosis, borderline personality disorders, schizophrenia, bipolar disorder, anxiety, major depression, intellectual disability, Alzheimer's disease, autism, ADHD, and addictions) were excluded. This increased the probability that the genetic profile that we examined was not confounded due to high comorbidity of aggression with other psychiatric disorders. Articles referring to aggression from the perspective of victimization and bullying were excluded. The publications were reviewed independently by two authors (V.V.O and P.J.R.), and when in doubt other co-authors were consulted until consensus on inclusion was reached.

Selection procedure and analyses

The search on review/meta-analyses resulted in 1 713 records (**Figure 1**). Duplicate entries were removed (*N*=27). Next, 1 660 records were excluded based on screening the titles and abstracts. In total, 26 potentially relevant reviews were retrieved for a full-text screening. Studies that did not fulfil or only partially fulfilled our criteria were excluded from the analysis (*N*=12), leading to the inclusion of 14 articles. Four additional reviews were added through the automated selection, leading to a total of 18 articles – 13 targeted and 5 systematic reviews. These were organized into the following categories: review type (targeted or systematic), definition of aggression, type of reviewed studies (heritability, candidate gene, GWAS), population (children, adolescents, adults), quantity and period of the publications included in the reviews (parameters are made on the basis of reference lists with inclusion of publications on the aggression-related traits), described genes and main conclusions.

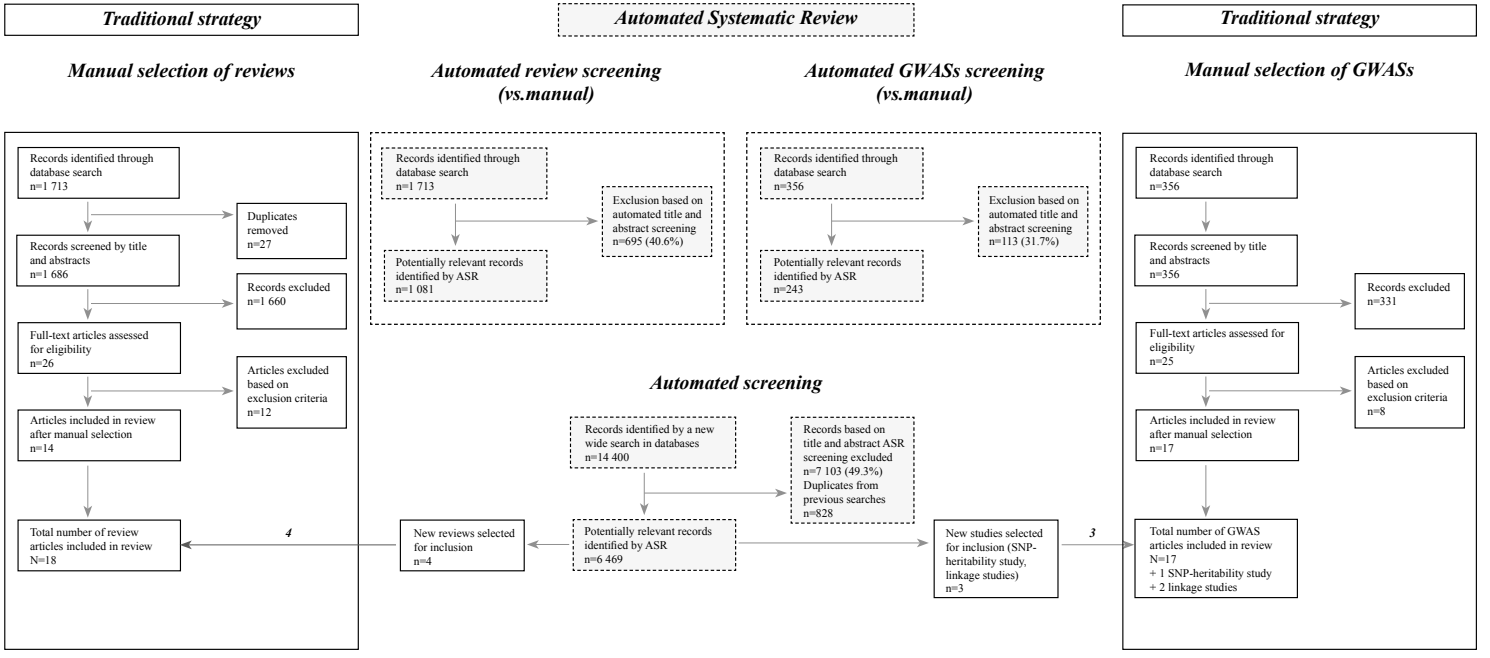


Figure 1. Flow diagram of literature selection

TABLE 1.
Inclusion and exclusion criteria for the systematic review

Selection criteria	Inclusion criteria	Exclusion criteria
Language	English	Non-English
Population	Human studies (all ages)	Animal studies
Use of term 'aggression'	Psychological/psychiatric	Disease characteristics (e.g. aggressive cancer, aggressive form of somatic diseases etc)
Psychiatric disorders	ODD, CD, ASPD	Victimization, victims of bullying Other neuropsychiatric and psychiatric disorders (e.g. psychosis, anxiety etc)
Discussion of genes	At least 2 genes associated with aggressiona	No genetic methods and information on genes associated with aggression

ASPD, antisocial personality disorder; CD, conduct disorders; ODD, oppositional defiant disorder
a This was done to exclude reviews focussing on a single candidate gene

The search for GWASs on aggression resulted in 356 records. A total of 331 were excluded based on screening of the titles and abstracts. This led to the retrieval of 25 potentially relevant studies for full-text screening. Studies that did not fulfil or only partially fulfilled our criteria were excluded (*N*=8), leading to the inclusion of 17 GWAS articles. Three additional studies were selected from the automated selection, including one SNP-heritability and two linkage studies. The studies were analysed by phenotype, sample characteristics, SNPs or genetic variants associated with aggression-related traits at $P < 1x10^{-05}$, genetic variants position in genes and chromosomes.

Several GWAS papers report findings on multiple (stratified) GWASs. Tielbeek et al.[27] adjusted for the fact that they performed three genome-wide association meta-analyses (GWAMA) by setting the genome-wide significance threshold at $P=1.67x10^{-08}$, whereas others did not apply such a correction. This threshold might be overly conservative as the GWAMAs are stratified, which makes the *P*-values nonindependent across GWAMA. Therefore, we maintained a significance threshold of $P=5.0x10^{-08}$ for all studies, and denote any SNP with a *P*-value below this threshold as genome-wide significant. While the traditional threshold might be too lenient in this context, we note that, when discussing GWASs, the *P*-value of a SNP in any given study is of less relevance than replication across GWASs.

Automated titles and abstracts screening

In parallel with the manual selection of titles and abstracts, another selection was made with the use of an automated selection tool 'Automated Systematic Review' (ASR) – software hosted at https://github.com[28]. This software allows for automated in- and exclusion of articles for systematic reviews based on the titles and abstracts of articles. This enabled a comparison between 'traditional' manual selection and the automated screening on performance characteristics (e.g. time spent on selection, false negative results). Furthermore, an additional selection was performed with the ASR on a large dataset of references to retrieve any new additional papers to our review, which would have been missed in the traditional search strategy (**Supplement S3**).

We trained a model using ASR. To do so, the model requires a training set based on expert knowledge, consisting of papers that are either labelled relevant or non-relevant (labels 1 = included, 0 = not) (see **Supplement S3: Figure S3.1**). To study the operating characteristics of the ASR, we used a dataset (*N*=2955) consisting of relevant and nonrelevant articles on the genetics of human aggression, as labelled by researchers. From this labelled dataset (*N*=2955), 500 records were repeatedly drawn at random as training sets. The number of relevant records in the training sets varied between 10 and 80 (e.g. 10 relevant records vs. 490 nonrelevant records), in increments of 10. These sets were used to train models to include relevant records and exclude nonrelevant records. For each model, we computed receiver operating characteristic parameters that were then used to select the optimal model (see **Supplement S1: Table S3.1, Figure S3.2**). We selected the model that returned the lowest false positive rate (FPR) while allowing for a maximum false negative rate of FNR = 0.03 at most. Note that FNR = 0.03 corresponds with a true positive rate of TPR = 0.97.

We applied the optimal model to predict classification in different searches: (1) reviews of genetics of human aggression (1 713 records); (2) GWASs on human aggression (356 records); (3) searches 1 and 2 combined (2 069 records) to analyse parameters of automated selection in comparison to manual selection.

Training sets were provided to the ASR for the reviews on aggression (26 relevant records out of 1713 [1.5%]) and the GWASs on aggression (25 relevant records out of 356 [7.0%]) (see **Supplement S3: Table S3.2**). The automated selection predicted 1 018 records out of 1 713 (59.4%) as relevant for reviews (including all prelabelled positives: TPR = 1.0; FPR = 0.59) and 243 records out of 356 (68.3%) for GWAS (including 24 prelabelled positives: TPR = 0.96; FPR = 0.66). Automated selection predicted 1 261 records out of 2 069 (60.9%) as important (including 50 prelabelled positives: TPR = 0.98; FPR = 0.60). The workload for manual selection was ~60 hours. This means that for the applied model and these set(s), the reduction in workload is expected to be ~23.5 hours. By allowing for a higher FNR in model selection, the workload could be reduced even further, although at the expense of missing more true positives.

Our automated selection repeated the traditional manual search with inclusion rates (100% for reviews [58.8% false positives], 96.0% for GWASs [66.2% false positives], 98.4% for reviews and GWASs combined [60.0% false positives]), 0 cases were false negatives for reviews, 1 case for GWASs, and 1 case for reviews and GWASs combined.

A new search on 'human aggression genes' was performed in the same databases without additional search terms and time limitation (14400 records) to detect new contributions to the systematic review, resulting in 55.8% included records. Exclusion of duplicate records resulted in 6469 records. From these, four reviews were added to the overview of reviews on aggression, and one SNP-heritability and two linkage studies were added to the GWASs review as additional information for the interpretation of GWAS findings. These seven studies were detected only by the ASR approach and did not appear in the traditional approach.

Results

We included 18 reviews on the genetics of human aggression in our analyses, each covering different periods and including varying numbers of studies (**Table 2**). The reviews cover more than 2000 studies on aggression.

What is considered to be aggression?

Reviews indicate that the phenotypic definitions of aggression vary considerably, and heterogeneity of the phenotypic definition is mentioned as a major hurdle in aggression research by multiple articles. Definitions of aggression, as well as the focal points of reviews, range from broadly defined externalizing and ASBs (see **Box 1**), which also include potentially nonaggressive behaviours like rule-breaking behaviour[29], to a narrow focus on chronic physical aggression[30]. Other reviews and studies focus more explicitly on psychiatric classifications like ODD, CD, and ASPD, which encompass aggressive acts and are correlated to ASB[2,31]. One review incorporated the analysis of genetics of aggression in suicidal behaviour[32]. Classifications, which are useful in clinical practice, tend to consist of constellations of heterogeneous ASBs (e.g. 'often initiates physical fights' vs. 'is often truant from school') and personality characteristics (e.g. 'having difficulty sustaining long-term relationships' vs. 'lacks concern, regret or remorse about other people's distress'[33]). Recent GWAS have also included broader definitions of risk tolerance and risky behaviors and externalizing behaviors[34,35].

Several authors proposed a focus on more homogeneous or dimensional constructs of aggression[2,29,30]. A dimensional construct is in line with the conceptualization that pathological aggression is situated on the extreme ends of a normal distribution[2]. Some authors see a risk in the dimensional approach and note that findings might become predominantly driven by variations within normal, adaptive levels of aggression[36]. However, if pathological levels of aggression are indeed the extreme end of a continuous phenotype, the same genetic and environmental factors should apply to both the normal range and extremes of the distribution.

In the end, concerns regarding heterogeneity and the impact of different phenotype definitions are empirical questions, which are currently also being asked in other GWASs of psychiatric disorders such as depression[37]. Such questions can be resolved, once well-powered GWASs are available, by estimation of genetic correlations among different phenotype definitions of aggression and can also be addressed through genetic modelling of twin and family data. For example, Hendriks et al.[38] analysed twin data collected by multiple instruments, commonly employed to measure aggression in children. While phenotypic correlations between different aggression scales could be low, a genetic multivariate analysis of these data showed high genetic correlations among different instruments. Such observations mean that different instrument tap into the same genetic liability and could be analysed simultaneously in GWAS.

Reviews that propose some sort of differentiation among aggressive behaviours often return to a distinction between reactive and proactive aggression. Reactive aggression is commonly described as impulsive and defensive, while proactive aggression is considered predatory and premeditated. Both types of aggression may involve similar

TABLE 2.
Reviews on genetics of human aggression

Review	Type of studies included	N papers with trait-related studies	Taxonomy of aggressive behavior (phenotype)	Samples
Baud, 2005	Heritability studies, CGS	91	Limited discussion of genetics studies of aggression, impulsivity and anger-related traits in suicidal behavior	Humans
Moffitt, 2005	Heritability studies (twins, adoption, family)	117	Antisocial behavior	Humans (children, adolescents, adults)
Craig and Halton, 2009	Heritability studies, CGS, GWAS	117	Human aggressive behavior; instrumental (proactive) and reactive	Humans
Tuvblad and Baker, 2011	Heritability studies (twin and adoption studies), CGS	138	Human aggressive behavior	Humans (children, adolescents, adults)
Anholt and Mackay, 2012	CGS, GWAS	127	Aggression as quantitative trait, pathological aggression (in substance abuse, psychiatric disorders, Alzheimer), externalizing behavior	Humans and animals
Vassos et al., 2014	CGS	185	Aggression and violence (categorical and continuous outcomes)	General population and specific subgroups
Provencal et al.., 2015	Heritability studies, CGS, GWAS, EWAS	176	Chronic physical aggression	Humans and animals
Zhang-James and Faraone, 2016	CGS	524 OMIM records	Aggressive and antisocial behavior, CD	Humans
Fernandez-Castillo and Cormand, 2016	CGS, GWAS, pathways and functions	198	Aggressive behaviors including aggression traits (aggressiveness, impulsive aggression, anger, externa-lizing behavior, violence, delinquency or criminality) or diagnostic categories (OD, CD, ASPD, CU, and psychopathy)	Humans
Veroude et al., 2016	Heritability studies, animal models, CGS, GWAS	378	RDoC: frustrative non-reward, defensive and offensive (or proactive) aggression. ODD, CD, APD	Humans (children, adolescents, adults) and animals
Waltes et al., 2016	Heritability studies, animal models, CGS, GWAS, EWAS	248	Human aggressive behavior, reactive (impulsive) and proactive (pre-mediated) aggression	Humans

Table 2. (continued)

Review	Type of studies included	N papers with trait-related studies	Taxonomy of aggressive behavior (phenotype)	Samples
Manchia and Fanos, 2017	CGS, GWAS, epigenetic, metabolomic, microbiome association studies	87	Aggression in mental illness	Humans
Beaver et al., 2018	Heritability studies, CGS, GWAS	40	Antisocial behavior, aggression, violence	Humans
Tremblay et al., 2018	Heritability studies (twin studies, adoption studies), CGS, epigenetic studies	123	Physical aggression	Human (children, adolescents) and animals
Davydova et al., 2018	Heritability studies, CGS, GWAS	78	Aggressive behaviour	Humans (children and adults)
Salvatore and Dick, 2018	Heritability studies, CGS, linkage, GWAS, GxE studies, rGE studies, epigenetics	96	Conduct disorder	Humans
Gard et al., 2018	Heritability studies, CTG, GWAS (meta-analyses)	56	Antisocial behavior, including aggression, violence and rule-breaking	Humans
Gescher, 2018	Epigenetic studies, EWAS	16	Antisocial traits, aggression, impulsiveness	Humans
Barr and Dick, 2020	GWAS, PRS	82	Externalizing behaviors/disorders such as alcohol or substance misuse, antisocial behaviors, aggression, and risk taking	Humans

CGS, candidate gene studies; GWAS, genome-wide association study; EWAS, epigenome-wide association study; PRS, polygenic risk score; GxE, genome-environment interaction; rGE, genome-environment correlation; OMIM, Online Mendelian Inheritance in Man; RDoC, research framework for investigating mental disorders; ODD, oppositional defiant disorder; CD, conduct disorder; CU, callous-unemotional traits; ASPD, antisocial personality disorder; APD, antisocial personality disorder.
Note: Updated by Odintsova et al, 2022 (in press) with addition of Gescher, 2018, and Barr and Dick, 2020.

biological systems. The aminergic systems (e.g. serotonergic, dopaminergic) have been proposed as likely to regulate both forms of aggression[39]. Interestingly, Runions and colleagues[40] argue that researchers studying reactive and proactive forms of aggression have conflated motivation (aversive vs. appetitive) and implementation (impulsive vs. premeditated) and propose that predatory aggression can also be impulsive in nature, defined as recreation instead of rage, while reactive aggression could also be delivered after a longer period of time, referring to reward instead of revenge.

The developmental aspect of aggression is a major theme in reviews[1,2,39,41–43]. Age of onset is often mentioned as an important differentiating factor for subtypes of antisocial behaviour, with aggression usually already present in early childhood, while rule-breaking behaviour and delinquency usually develop during adolescence. Tremblay[44] proposes a developmental framework of aggression among a covert/overt axis and a second destructive/nondestructive axis as the most viable constructs to subtype disruptive behaviour (aggression, opposition-defiance, rule breaking, and stealing-vandalism). Children who display destructive and overt disruptive behaviours, especially those exhibiting chronic physical aggression, experience more risk factors early in life, engage in aggression from a young age, and have a more persistent developmental course of aggression and ASB. A differentiation on age of onset is considered especially relevant in reviews which include epigenetics. Epigenetic changes may be triggered by early life adversity[30,42,45,46], although variation in epigenetic marks can also reflect influences of DNA polymorphisms[47].

In research, aggressive behaviour often is measured by questionnaires, such as the Achenbach System of Empirically Based Assessment scales (ASEBA[48], the Strengths and Difficulties Questionnaire (SDQ[49], or the Buss Durkee Hostility Inventory (BDHI[50]). Aggression scales in such instruments may include items that reflect behaviour that is related to aggression, but would not be considered aggression based on item content. For example, the ASEBA Aggressive Behaviour scale of for children contains items like ‘Argues a lot’ or ‘Gets in many fights’, but also ‘Unusually loud’ or ‘Suspicious’. Measures can also derive from observational studies, especially in younger children, and some experimental paradigms are available to measure aggression in across wider age ranges. Such experiments can, however, not cover the full spectrum of aggressive behaviour and, perhaps even more critically, cannot be applied in epidemiological samples.

There is a divergence between measurement of aggression in research projects compared to how (pathological) aggression is defined in clinical practice. Questionnaires are used as tools by clinicians, but the presence of these behaviours is mostly determined by interviews with the patient, and others who know the person (e.g., parents, teachers), by observation, and by the patient’s (criminal) records. Psychiatric disorders that include aggressive behaviours or disorders, which are correlated to aggressive and antisocial lifestyles, are dependent on classification systems like the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Classification of Diseases (ICD). In these classifications, a dichotomy is applied in which a disorder is either present or absent, largely ignoring the dimensional nature of human behaviour. In genetic studies, a focus on the dichotomy rather than on continuous variation, may lead to a loss of statistical power[51].

Another important question, especially in clinical settings, is when aggression becomes pathological. Some aggressive behaviours are clearly defined as pathological, like aggressive behaviours that define Conduct Disorder (e.g., 'Has used a weapon that can cause serious physical harm to others'), or Antisocial Personality Disorder (e.g., 'Irritability and aggressiveness, as indicated by repeated physical fights or assaults'). In contrast, other aggressive behaviours are less clearly considered pathological, because they occur to some extent in all individuals, like anger or hostility. This even is the case for some aggressive behaviours which are part of disruptive behaviour disorders (e.g., ODD: 'often argues with authority figures'). For aggression to be pathological, it is essential that aggressive behaviours cause clinically significant impairment in social, academic, or occupational functioning.

Approaches in genetics of aggression studies and the current status quo

There are several designs to study the genetic aetiology of aggression, with the two major ones being genetic epidemiological/behavioural genetic approaches on the one hand and molecular genetic approaches on the other (**Figure 2**). Behavioural genetic studies have a long and successful history[52]. More recently, molecular genetic studies have seen enormous breakthroughs with the development of techniques like GWASs[26,53].

Behavioural genetic approaches

Numerous studies focused on explaining the aetiology of aggression and ASB through family, twin, and adoption studies, which can disentangle genetic and environmental influences. Twin models enable researchers to divide the variance for a trait, or the liability to a disorder, into genetic and non-genetic components. The genetic variance component often is defined as the additive (A) effects of many genes. Environmental variance components consist of environmental influences common to siblings from the same family (C), creating resemblance of family members through environment rather than through genetics, and a unique or non-shared environmental component (E). Unique environmental influences affect family members in different ways[54]. Unsystematic influences such as measurement error also are included in the E component, unless explicitly modelled. In general, reviews indicate that additive genetic factors explain around 50% of the variability of aggressive behaviour[1,29,55,56]. The estimate varies around 50% across studies, with some reviews reporting somewhat higher heritability estimates (65%) and others giving estimates for aggression and ASB that vary more (e.g., 38%-88%[2]; 28%-78%[1]). Physical aggression seems to show larger heritability estimates (65%) than reactive (20-43%) and proactive aggression (32-48%), while rule-breaking behaviour, which is often aggregated with aggression indices, also shows a heritability around 50%[39,57]. Heritability estimates of aggressive behaviour were higher in children with stable callous unemotional traits (81%) compared to children low in callous unemotional traits (30%)[57]. This suggests a larger influence of genes on children with more severe aggressive tendencies[57]. Contributions of shared environment are relatively small and decrease with age, with the vast majority of adult studies not reporting any shared environmental influences[1,2,39]. Thus, research in behaviour genetics clearly indicates that there is a substantial genetic component to aggressive behaviour in humans. In longitudinal studies, heritability estimates of aggression and ASB increase somewhat from childhood through adulthood[1,2,39]. Genetic factors also contribute to the stability of

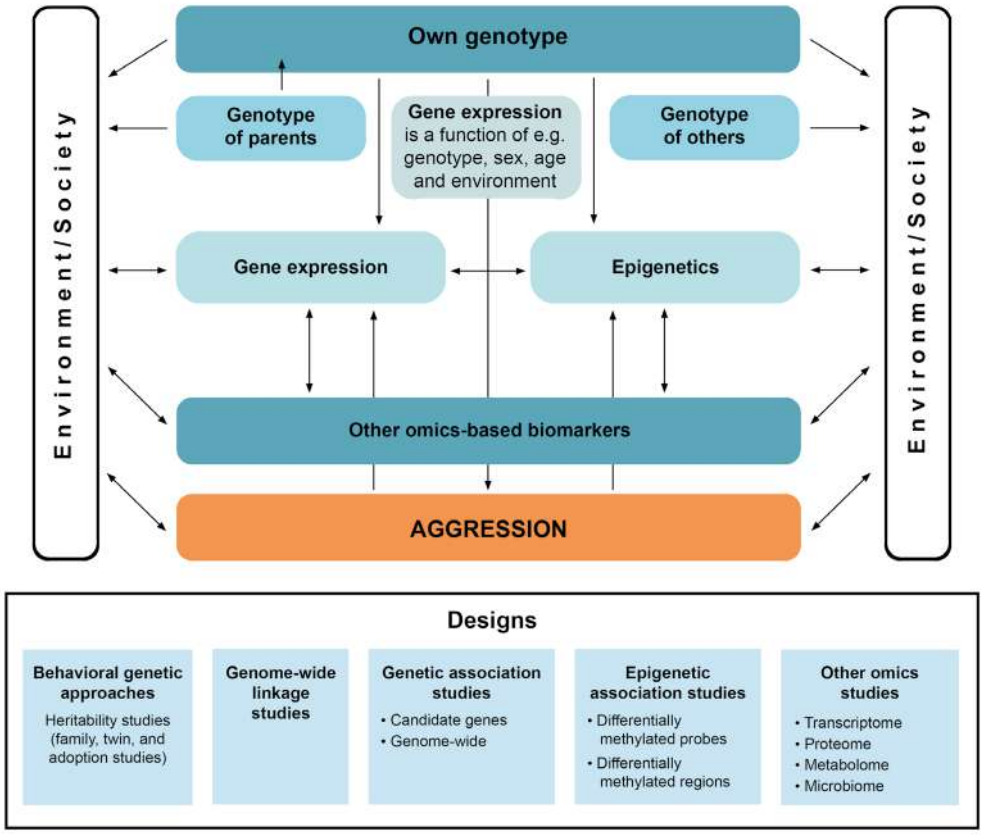


Figure 2. Behavioral genetic and molecular studies of aggression: interplay of genetic, epigenetic, other omics, and environmental factors in aggressive behavior

Updated by Odintsova et al, 2022 (in press).

aggressive behaviour during preschool and school age, and puberty[39,58]. Measurement instrument, and also rater, seem to influence heritability estimates, with heritability based on parent-report and teacher-report estimated as higher than those based on self-report and observational studies. Studies based on self-report tend not to find any shared environmental influences[1], but such studies are not available for younger children. Unlike parent or teacher reports, observational studies more often give an assessment of aggression at one particular moment in time only. Parent- and teacher-reports tend to provide phenotype information that is more averaged over longer periods of time and are similar in terms of heritability estimates. Parent-report leads to higher estimates of shared environmental influences than teacher-report, when parental characteristics that influence ratings of multiple children (e.g. twins or siblings) are not taken into account. When twins have different teachers, similarities between them tend to decrease. This may reflect actual differences in aggressive behaviour with different teachers and/or

different settings, but may also reflect teacher characteristics that influence assessments of multiple children.

In summary, heritability is estimated consistently around 50%, with some variation that may be due to different conceptualization of aggressive and ASBs, with more severe types of aggression showing higher heritability.

Heritability estimates of aggression and antisocial behaviour may differ between environments suggesting an interaction between genes and environment (GxE). Proposed putative environmental moderators are familial adversity (e.g. maltreatment and parental delinquency), social disadvantage (e.g., poverty and bad neighbourhoods), violent media exposure, and alcohol use. Tuvblad and Baker[1] argue that, compared to genetic factors, environmental influences are relatively more pronounced for antisocial behaviours in the presence of high environmental risk and disadvantaged environments. Conversely, genetic influences will be more pronounced when environmental risk factors are absent or less prominent. In one study, the moderating effects of neighbourhood seemed to be specific to the heritability of nonaggressive antisocial behaviour, while heritability estimates of aggressive antisocial behaviour were not influenced by neighbourhood disadvantage[59]. Such findings underscore the differential influence of environmental adversity on certain types of ASB, with aggressive behaviour showing less sensitivity to environmental influences than other types of ASB. Later reviews, however, indicate mixed findings. Some reported an increase in genetic variance in the presence of environmental risk. To illustrate, when young children were subjected to high levels of maternal disengagement, genetic factors explained more variance in later conduct problems[39,60]. An increase in heritability of externalizing disorders was also found when young adults were exposed to a combination of risk factors (e.g. antisocial or lack of prosocial peers, relationship problems with parents[2,61]).

Depending on the type of aggression, mean levels of aggression often are higher in males than in females. Differences in heritability estimates, however, between males and females are modest or absent. According to Tuvblad[1], heritability did not differ significantly between genders across different twin studies, either quantitatively or qualitatively (see also Vink et al.[62]). These studies mainly included mother-reports of childhood aggression and heritability estimates were higher in males than in females when self-report data were analysed[39]. It has been suggested that gender differences in heritability become more pronounced from adolescence, which could be indicative of the 'Young Male Syndrome', in which the onset of puberty and increasing levels of testosterone are related to increases in aggression in 12- to 25-year-old males[55]. This would also suggest a possible role of genes related to androgen synthesis and function in the development of aggression from puberty onwards.

In summary, twin studies highlight the importance of genetic influences, with estimates of the heritability of aggression and antisocial behaviour often reported to be around 50%[41], without much evidence for sex differences in heritability estimates. Such significant heritability is a first requirement for initiating studies that aim to find molecular signatures in the DNA sequence that are associated or causally related to the phenotype.

TABLE 3.
Overview of genome-wide suggestive and significant associations with aggression-related traits at P≤1x10^{-05} per GWAS

Study	Phenotype	Sample	$N_{variants}$	N_{genes}	Genes	Summary of main findings
Sonuga-Barke et al., 2008	CD using PACS	*N*=909 probands in trio. Males ~87%. ADHD diagnosis ~99%. Age range: 5-17 years. European Caucasian ancestry	18	7	GxE interaction with "mother's criticism" PPM1K, ZBTB16 GxE interaction with "mother's warmth" *RBFOX1(A2BP1), ADH1C (proximal), MFHAS1, SLC6A1, RIT1 (proximal)*	Suggestive GxE interactions were reported for 18 SNPs, of which 3 SNPs also showed a suggestive main effect. For both the main and interaction effects, no SNP reached genome-wide significance
Anney et al., 2008	CD using DSM-IV criteria for CD, PACS and CPRS-R:L, gathered the symptom on a less severe behavioral characteristic of an oppositional defiant individual.	*N*=938 probands in trios. Males ~87%. ADHD diagnosis ~99%. Age range: 5-17 years. European Caucasians ancestry	54	41	*LIG4 (proximal), ABHD13 (proximal); AMOLT1 (proximal), CWD15 (proximal), KDM4D (JMJD2D) (proximal); FLJ16077; RXFP1 (proximal); PAWR; LOC729257; SPATA8 (proximal); YWHAZ (proximal); FLJ31818, GPR85 (proximal); KIRREL3; PRPRD (proximal); ATP8B1 (proximal); MYRFL (c12orf28); LIG4 (proximal), ABHD13 (proximal); PKD1L2; c16orf46 (proximal); PKD1L3; KIAA0174 (proximal),; DHODH (proximal); c5orf16 (proximal); c5orf15 (proximal); FLJ39064; FZD10 (proximal); FLJ39063; FZD9 (proximal); FLJ39062; FZD8 (proximal); ILVBL(FLJ39061); FZD7 (proximal); ETV3L (proximal), ETV3 (proximal); FLJ17340; GSX1 (proximal), PDX1 (proximal); PITRM1 (proximal); RBFOX1(A2BP1); GLT25D2 (proximal); RGL1*	Suggestive associations were reported for 54 SNPs. These SNPs were located in 13 genes and/ or were within a 200kb window of 28 additional genes. The top five association signals were observed on Chr 13, 21, 11, 4, and 12
Viding et al., 2010	ASB/CU: Teacher-rated conduct problems and CU traits using SDQ; 3-point scale	*N*=600 from twin cohort (high- and low-scoring of AB) Males 69%. Age=7 years. Caucasian ancestry.Replication *N*=586. Males 71%.	0	0	Suggestive in replication (*P*=4,77E^{-05}) *KCNMA1*	In both the discovery and replication study, no SNP reached genome-wide significance. Several top SNPs were located near neurodevelopmental genes such as *ROBO2* (*P*=4.61x10^{-03})

Table 3. (continued)

Study	Phenotype	Sample	$N_{variants}$	N_{genes}	Genes	Summary of main findings
Dick et al. 2011	CD: retrospective report of DSM-IV CD symptoms, natural log as primary CD measure	*N*=3,963 (N_{cases}=872, $N_{controls}$=3091). Age range: 18-77 years. European Americans, African Americans	29	10	In a sample with mixed ancestry *C1QTNF7*; PDE10A; SELPLG; TOX2; LOC343052; ERCC4*	European sample: were only reported for the top 20 SNPs that came out as suggestive/significant for the mixed analysis. None of the SNPs were suggestively associated with either phenotype within the European sample. Mixed sample with European and African ancestry: 4 SNPs reached genome-wide significance level for CD_{symp} – but not for CD_{cc} – two of which were located inside *C1QTNF7*. The other two significant SNPs were not located near any gene
Merjonen et al.., 2011	Anger in hostility dimensions measured by the Irritability Scale of the Buss-Durkee Hostility Inventory in four time points over a 15-year interval	*N*=2443. Males 46%. Age range: 15-30 years. Followed up for 15 years. European Caucasians ancestry (Finnish population)	20	2	*SHISA6; PURG*	One SNP reached significance $P<9x10^{-8}$: Chr 17: rs11656526, closest gene *SHISA6*. Many associations with anger approached significance, among them SNPs located close to genes *PURG*

Table 3. (continued)

Study	Phenotype	Sample	$N_{variants}$	N_{genes}	Genes	Summary of main findings
Tielbeek et al., 2012	ASB according to DSM-IV for CD Cohort 1: non-diagnostic measure covering seven items related to antisocial behavior, case status was 3 symptoms or more Cohort 2: Diagnostic measure of ASPD, cases had a diagnosis of ASPD except for criterion D (the occurrence of antisocial behavior is not exclusively during the course of schizophrenia or a manic episode)	Combined sample *N*=4816 (298 cases, 4518 controls). Males 41%. Age range cases: 18-74 years. Age range controls: 18-77 years. Australians	22	12	*DYRK1A; AL590874.1; CIB1; SEMA4B; TTC7B; IMMT; CSMD1; REEP1; RP11; BAZ2B; STK32A; VRK1*	Only results for top 50 SNPs were reported. No SNP reached genome-wide significance, but 9 were suggestively associated with DP. Out of these 9, 7 were located within 4 genes. Suggestive evidence for developmentally expressed genes operant in hippocampal dependent memory and learning associated with CBCL-DP is found
McGue et al., 2013	Behavioral Disinhibition; composite score consisting of five symptom counts for CD, ASB, Dissocial behavior, Delinquent Behavior Inventory, Aggressive underscore	*N*=7,188. Males 46%. Age: adults. Caucasian ancestry	4	1	*GLIS1*	Genome-wide suggestive levels were reached by 4 SNPs, tagging 1 gene

Table 3. (continued)

Study	Phenotype	Sample	$N_{variants}$	N_{genes}	Genes	Summary of main findings
Tiihonen et al., 2015	Violent offending; at least one sentence for violent offence. Extreme violent offending; 10 or more sever violent crimes	Violent offending: N_{cases}= 360. Males 94%. Extreme violent offending: N_{cases}= 56. Males 97% $N_{controls}$=5983. Males 57% Age: mean(s.d.) = 29.4(8.2). Finnish population	14	9	Violent behavior: *SPIN1; NTM; ATP10B (proximal); PRMD2 (proximal); PLCB1; NXPH1 (proximal)* Extremely violent behavior: *CDH13; PRUNE2; LOC101928923*	Genome-wide suggestive levels for violent behavior were reached by 10 SNPs, mapping to 6 genes. Additionally, 4 suggestive SNPs (3 genes) were reported for extreme violent behavior
Mick et al., 2014	Angry temperament and angry reaction measured by SSTAS	N = 8,747 from Atherosclerosis Risk in Communities Study. Males 47%. Age range: 45-64 years European ancestry	8	5	Angry temperament: *FYN (proximal), IYD (proximal), ZNFX1 (proximal), STAU1 (proximal), DDX27 (proximal)* Angry reaction *(P<6 x10-03): PHEX (proximal), SLC39A8 (proximal), MBOAT1(proximal), PLEK (proximal)*	*P*-values results from phenotypes adjusted for principal components representing genetic structure were used. Four SNPs reached suggestive levels of significance for angry temperament. Five SNPs reached suggestive levels for angry reaction $P<6x10^{-03}$, tagging four genes. Both scales were also dichotomized and treated as case-control phenotype, for which no SNP returned suggestive results
Pappa et al., 2016	Predominantly parent-reported child aggressive behavior. SDQ, CBCL, and other (parent rated questionnaires) in different cohort	*N*=18,988; 9 cohorts. Age range: 3-15 years. North European ancestry	76	16	Overall: *LRRTM4 (proximal)*; PDSS2; TRIM27 (proximal); MRC1; MECOM; CASC17 (proximal)* Early childhood: *COL13A1; SDK1 (proximal); LOC101928923; TSG1 (proximal); LOC727982 (proximal)* Middle childhood/early adolescence: *LRRTM4 (proximal); LOC101927797 (proximal); OPCML; COL13A1; GRIA1; ASBA; CNTN4*	Meta-analysis of nine cohorts reported one genome-wide significant hit. 35 SNPs reached suggestive levels for the overall GWAMA. These SNPs are located inside three genes and near three others. 10 and 31 SNPs reached suggestive levels for GWAMA on early and middle childhood/early adolescence AGG, respectively. Some of these SNPs overlap with the top hits reported in the overall GWAMA. In total suggestive associations were reported for 76 SNPs (66 unique) located in or around 16 genes

Table 3. (continued)

Study	Phenotype	Sample	$N_{variants}$	N_{genes}	Genes	Summary of main findings
Rautiai-nen et al., 2016	ASPD (violent criminals, substance abuse, maltreatment). ASPD diagnoses, SCID-II items for DSM-IV	*N*=6,220 (Cases 370, Controls 5850). Males 59%. Age: mean (s.d.) $_{ASPD}$ = 34.5(8.0); mean (s.d.) $_{Controls}$ = 55.0(13.2). Replication N = 3939(Cases 173, Controls 3766). Males 43%. Age: meand (s.d.) $_{ASPD}$ = 34.2±9.2; mean (s.d.) $_{controls}$ = 55.0±17.0. Finnish population	6	1	Cross-sex: *LINC00951 (proximal)** Males only: *LINC00951 (proximal)*	Results based on meta-analysis across discovery and replication reported that for the cross-sex GWAMA, 1 SNP reached genome-wide significance while another SNP ~10Kbp away reached suggestive levels. The closest gene to these SNPs is *LINC00951*. In the male-specific GWAMA, four SNPs reached suggestive levels, two of which are the same ones as the SNPs reported in the overall GWAMA. The other two SNPs are within ~50Kbp
Aebi et al., 2016	ODD. CPRS-R: L. Continuous: defiant/ vindicative; irritable Case-control: low/ moderate OPP vs irritable /severe OPP	*N*=750 with available ODD. Males 87.8%. Age range: 5-18 years. European Caucasian ancestry	53	14	*ADAM12; MYLK2 (proximal); OR2AG1 (proximal), OR2AG2 (proximal); BCL2L1; TPX2; DDX24 (proximal), ASB2 (proximal); RARB; RUNX1T1; FOXS1 (proximal); TTLL9 (proximal); COX4I2; SOX5; MYLK2*	Results based on multivariate GWAS only reported that 53 SNPs reached genome-wide suggestive levels, which are located inside and/or near 14 unique genes
Brevik et al., 2016	Childhood aggressiveness in adult ADHD Adult sample: retroactive measure of childhood symptoms of ADHD. Child sample: CPRS-R:L, subdivided in defiant/vindictive and irritable dimension	N adults=1060 patients with ADHD. N children= 750 with ADHD. European Caucasian ancestry	65	20	*NTM; CSMD1; KRT18P42 (proximal); TEPP; CPNE4; MICAL2 (proximal); LOC101929236; LOC101927464; NR_110053.1; H3F3A; LOC105370057; ACBD3 (proximal); LOC101929156; LOC105376469 (proximal); LOC105373223 (proximal); SPINK2; PHLPP1; UFM1*	Results based on meta-analysis across adult and children samples reported that 65 SNPs – located in or near 20 genes – reached suggestive levels of associations. The strongest signal was observed at rs10826548 on Chr 10 located within the transcript of a long noncoding RNA (P=1.07x10^{-06}), closely followed by rs35974940 in NTM (P=1.26x10^{-06})

Table 3. (continued)

Study	Phenotype	Sample	$N_{variants}$	N_{genes}	Genes	Summary of main findings
Tielbeek et al., 2017	Broad-spectrum ASB. Development and well-being assessment, conduct disorder scale, count of the number of APD criteria, rule-breaking behavior, Teacher report Form, Antisocial Process Screening Device, Retrospective CD, SCID-II for DSM-IV disorders, CBCL: conduct problems (reported by mother), DSM-IV CD criteria	*N*=16,400. Males 47%. Replication N=9,381. Mean age range across cohorts: 6.7-56.1 years. Ancestry: Mixed	80			GWAMA across five cohorts. Only independent signals are reported. The cross-sex GWAMA reports 20 suggestive associations, of which 2 are InDels. Two significant associations were found for the female-specific GWAMA. These two SNPs are located on Chr 1 and 11, respectively. The male-specific GWAMA returned one significant association on the X-chromosome. The female- and male-specific GWAMAs returned 37 and 20 suggestive associations, respectively. In total 80 unique variants (64 SNPs) were associated with ASB. ASB has potential heterogeneous genetic effects across sex
Montalvo-Ortiz et al., 2018b	Cannabis related physical aggression assessed with the question, "Did you ever get into physical	*N*=2,185 African Americans. Males 61%. *N*=1,362 European Americans. Males ~64%.	280**	43	European ancestry *LPPR1; ARHGEF3; RARB; TMEM92; ERBB4; CCDC171; ATP10A; UST; GPRC5B; CDH13; GRIN2B*	European-American sample: suggestive associations were found for 76 variants, of which 7 were structural variants. The 76 variants implicate 11 genes
	fights while using marijuana?"	Replication *N*=89 African Americans. Males 49%. Exposed to cannabis use. Age mean ~ 37-45			African ancestry *PSMD1*; HTR2B*; CCDC157; TBC1D10A; GSG1L; THSD7B; BRINP1; CNTN3; NSG2; SF3A1; SOD3; ADGRV1 (GPR98); KLHL3; SEC31A; ABR; TSPEAR; TMEM53; CCDC141; STAB2; RTN1; CDYL; UBE2H; LRMDA (C10orf11); ANO4; STRC; TASOR2 (FAM208B); SERTAD1; ARMH1 (C1orf228); CEP126 (KIAA1377); ABCA13; SLC17A6; LRRC4C*	African-American sample: the top SNPs included rs35750632 in *PSMD1* and rs17440378 in *HTR2B*. Based both on its demonstrated contribution to aggressive behavior and functional annotation analysis, *HTR2B* is suggested to be the relevant gene

Table 3. (continued)

Study	Phenotype	Sample	$N_{variants}$	N_{genes}	Genes	Summary of main findings
Ip et al., 2021	Childhood aggressive behavior measured by multiple assessors and multiple tools	*N*=328,935 observations from 87,485 individuals. Age range= 1.5 to 18 years. European ancestry	185	3	*ST3GAL3; PCDH7; IPO13*	Suggestive associations were found ($P< 10^{-5}$) for 185 SNPs. Gene-based analysis revealed three significant gene associations. All three genes have been previously associated with educational traits

From left to right, columns indicate (1) study, (2) phenotype description, (3) sample description, (4) number of (unique) associated SNPs/variants, (5) number of (unique) genes, (6) gene names, and (7) summary of main findings. Selection of associated with aggressive behavior genes presented in the table is done on the base of associated SNPs at $P<1x10^{-05}$ (nominally significant). Genes are sorted by ascending *P*-value for SNPs (the lowest level if gene is associated with several SNPs). When gene name has a new name in HUGO, the old name used in the study is given in brackets. The nearby location of nominally significant SNP is given in brackets (proximal), in other cases the location is intragenic.

* Genes for SNPs with genome-wide significance ($P<5.0x1^{0-08}$). ** For Montalvo-Ortiz et al. (2018b) SNPs, variants and genes are included at $P<1x10^{-06}$.

ASB, antisocial behavior; ASPD, antisocial personality disorder; BDHI, Buss-Durkee Hostility Inventory; CBCL, Child Behavioral Checklist; CD, conduct disorder; Chr, chromosome; CPRS-R: L, long version of the Conners Parent Rating Scale; CU, callous-unemotional traits; DP, dysregulation profile; DSM, Diagnostic and Statistical Manual of Mental Disorders; GWAMA, genome-wide association meta-analysis; GWS, genome-wide significant; NA, not available; ODD , oppositional defiant disorder; PACS, Parental Account of Childhood Symptoms; SCID, Structured Clinical Interview for Disorders; SDQ, Strengths and Difficulties Questionnaire; SSTAS, Spielberger State-Trait Anger Scale.

Note: Updated by Odintsova et al, 2022 (in press) with addition of Ip et al, 2021.

Integrating data on genetics of aggression from molecular genetic studies

Genetic linkage and candidate gene studies. Molecular genetic studies include genetic linkage and association studies, either genome-wide or with a focus on a limited number of candidate genes or candidate regions. In *linkage studies,* DNA markers are assessed in related individuals to investigate the inheritance of markers with known chromosomal locations together with aggression in pedigrees. Sometimes candidate regions to be investigated are suggested from studies in other species. With the arrival of large scale association studies, linkage studies, which require family-based designs, have become less common, but early studies have suggested regions on three chromosomes that could be associated with aggression. Dick et al.[63] analysed retrospectively reported childhood conduct disorder in an adult sample from COGA (Collaborative Study on the Genetics of Alcoholism). Regions on chromosomes 19 and 2 may contain genes associated with risk of CD. The same region on chromosome 2 has been linked do alcohol dependence in this sample. Criado et al.[64] in a linkage study of cortical even-related oscillations associated with ASPD and CD suggested that chromosome 1 may contain a genetic locus for ASPD/CD.

Genetic association studies initially were candidate gene studies. These require a priori knowledge of or hypotheses about which genes are implicated in the aetiology of the trait of interest. For aggression, associations were considered for genes from the serotoninergic (5-*HTTLPR* (5-hydroxytryptamine (serotonin) receptors), *SLC6A4* (solute carrier family 6 member 4), and dopaminergic (dopamine receptors genes *DRD4, DRD2, DRD5,* and *SLC6A3* (solute carrier family 6 member 3)) and GABAergic systems (e.g., genes that code GABA (gamma-aminobutyric acid) receptors, like GABRA2 (gamma-aminobutyric acid type A receptor alpha2 subunit)), as well as genes related to catecholamine catabolism (*MAOA* (monoamine oxidase A), *COMT* (catechol-O-methylt ransferase))[2,29,42,43,57]. Other studies focused on associations with the genes involved in stress response pathways[39,55]; hormone regulation (e.g., *AVPR1A* (argenine vasopressin receptor 1A))[2,29,39,65]; hypoglycaemia and insulin secretion[55]; and neuronal transcripts and brain expression patterns[8,39,55,57]. Candidate gene studies have been criticised[66], since it became clear that often the findings for candidate genes are often not replicated in well-powered GWAS[67,68]. It is likely that this also extends to studies of aggression, but the status of the candidate genes for aggression must await well-powered GWASs.

Many reviews agree that aggression is a polygenic trait influenced by many genes and that each explains a small proportion of the phenotypic differences. However, there may be an overlap between genes of large effect underlying monogenic disorders and those affecting continuous variability of related quantitative traits. Extending the idea of a shared genetic basis between Mendelian disorders and polygenic traits, one alternative approach based on the search for genes for aggression in studies of rare, functional genetic variants associated with aggression phenotypes catalogued in Online Mendelian Inheritance in Man (OMIM;[4]). Most of these genes had not been implicated in human aggression before, but the most significantly enriched pathways (e.g. serotonin and dopamine signalling) had been previously implicated in aggression. Among these genes, only two were previously related to aggression (*MAOA, GRIA3* (glutamate ionotropic receptor AMPS type subunit 3). New associations were found with genes (e.g. *CAMTA1*

(calmodulin binding transcription activator 1), *APBB2* (amyloid beta precursor protein binding family B member 2), *DISC1* (DISC1 scaffold protein), and others), which implicated in cell-to-cell signalling and interaction, nervous system development and function, and behaviour. The novel genes and pathways identified in this study suggested additional mechanisms underlying aggression.

Genome-wide association studies

Genome-wide association studies (GWASs) investigate millions of Single Nucleotide Polymorphisms (SNPs), under a continuous or dichotomous, case/control model. The result is a list that, for every variant, indicates the expected increase in a trait (continuous) or genetic liability (dichotomous) for every copy of an effect allele. Due to the large number of tests, the genome-wide significance level is set at P=5.0x10^{-08} [69], to properly control for the type I error rate. This adjusted threshold already considers the fact that neighbouring SNPs are not inherited independently from one another. However, the nonindependent inheritance of SNPs indicates that association tests between noncausal SNPs and the trait of interest contain a part of the polygenic signal[70]. As such, even when only a limited number of SNPs reach this stringent significance level, there is signal in the other association tests. The weighted effects of all the genetic variants involved in aggression could produce a polygenic risk score with a certain predictive value[71].

Many reviews discussed a whole-genome approach to understanding aggression, but only three have done so in a systematic manner[2,29,39]. We will summarize findings for genes harbouring, or in close proximity to, variants that reached genome-wide (P≤5.0x10^{-08}) or nominal (P≤1.0x10^{-05}) significance levels in all GWAS of aggression phenotypes to date. These include aggression-related phenotypes, i.e. anger, hostility dimensions, aggressive behaviour, physical aggression, ASB, violent offending, CD, ODD, and ASPD.

To provide a complete picture of the GWAS literature available, we chose to include phenotypes, which clearly include aggression, but are sometimes conflated with other ASBs (e.g. rule breaking) or personality characteristics (e.g. being suspiciousness and being loud). These phenotypes can be found in **Supplement S4**. Most GWASs on aggression were performed in child and adolescent samples of European ancestry, in which aggression was assessed using rating scales (**Table 3**).

GWAS studies have mainly resulted in nominal associations between genetic variants and aggression-related traits and disorders. Collectively, these studies reported 10 genome-wide significant findings[27,72–74]. Five of these variants are located inside or close to four genes: *LINC00951* (long intergenic non-protein coding RNA 951)[73], *C1QTNF7* (C1q tumor necrosis factor-related protein 7)[72], *PSMD1* (proteasome 26S subunit, non-ATPase 1), and *HTR2B* (5-hydroxytryptamine receptor 2B)[74]. Lastly, the five remaining significant SNPs are located on chromosomes 11[27,72], 13[72], 1 and X[27].

In a mixed sample of subjects from European and African-American ancestry, three SNPs inside *C1QTNF7* were significantly associated with CD symptoms in adults with substance dependence[72]. When the sample was split on the basis of ancestry, no SNPs reached suggestive levels in the European-American sample. In the African-American

sample one out of the three SNPs reached suggestive levels (minimum P=4.35x10^{-06}), along with two additional suggestive findings (minimum P=2.67x10^{-07}). *C1QTNF7* is less expressed in the brain, compared to such tissues as endometrium, gall bladder, lungs, ovaries and 18 other tissues, and has a potential role in maintaining energy balance[75].

In a study focusing on ASPD in Finnish criminal offenders, Rautiainen and colleagues[73] found one hit (rs4714329, P=1.6x10^{-09}) in the cross-sex meta-analysis. This variant is in close proximity to *LINC00951* (long intergenic non-protein coding RNA 951). The same SNPs returned suggestive associations in the male-specific GWAMA of ASPD (P=1.38x10^{-07}). The signal from these variants was specific for ASPD, and did not cover a broader range of criminal behaviour. Montalvo-Ortiz and colleagues 74 found that SNPs located in the *HTR2B* (P=2.16x10^{-08}) and *PSMD1* (P=1.79x10^{-08}) genes were significantly associated with cannabis-related physical aggression in African-Americans, but these SNPs did not reach even suggestive significance in European-Americans. Cannabis use has been associated with greater impulsive decision-making and increased aggressive behaviour. Notably this is the only GWAS study which focused purely on physical aggression.

Anney and colleagues[76] listed 54 SNPs nominally associated with conduct problems. These SNPs tagged 41 genes, three of which are with known functions and are involved in the regulation of dopamine receptor D2 signalling (*PAWR* (pro-apoptotic WT1 regulator)), synaptic plasticity (*KIRREL3* (kirre like nephrin family adhesion molecule 3)), and neuronal development (*RBFOX1* (ral guanine nucleotide dissociation stimulator like 1)). Sonuga-Barke and colleagues[77], analysed interactions between CD symptoms and maternal warmth. Nominal effects were found for SNPs located in genes involved in brain maturation, neurotransmission, neuronal development and regeneration. Viding and colleagues[78] examined teacher-reported conduct problems in children and found no suggestive SNPs (minimum P=4.6x10^{-05}).

For adult ASB[79], the strongest signal was for a SNP (rs346425; P=2.51x10^{-07}) located on chromosome 5. Salvatore and colleagues[80] in an adult ASB sample observed the strongest association for rs4728702 (P=5.77x10^{-07}), located in *ABCB1* (ATP binding cassette subfamily B member 1) on chromosome 7 that may confer general risk across a wide range of externalizing behaviours. Enrichment analyses further indicated involvement of immune-related pathways. Two GWASs compared cohorts of Finnish violent offenders to the general population[73,81], and obtained association signals at genes involved in neuronal development[81] and adaptive immunity[73].

Aebi and colleagues[82] hypothesized that *BCL2L1* (BCL2 like 1) is likely associated with oppositional behaviour, because of its influence on presynaptic plasticity through regulation of neurotransmitter release and retrieval of vesicles in neurons. Brevik and colleagues[83] applying gene-based tests observed *NTM* (neurotrimin) as the top gene, that is differentially expressed in aggression-related structures of the amygdala and the prefrontal cortex in early stages of brain development.

Merjonen and colleagues[84] saw suggestive associations for SNPs that lie inside genes involved in the maintenance of high frequency synaptic transmission at hippocampal synapses, and regulating synaptic activation (*SHISA6* (shisa family member 6)) in a Finnish population sample. Mick and colleagues[85] found associations for SNPs that

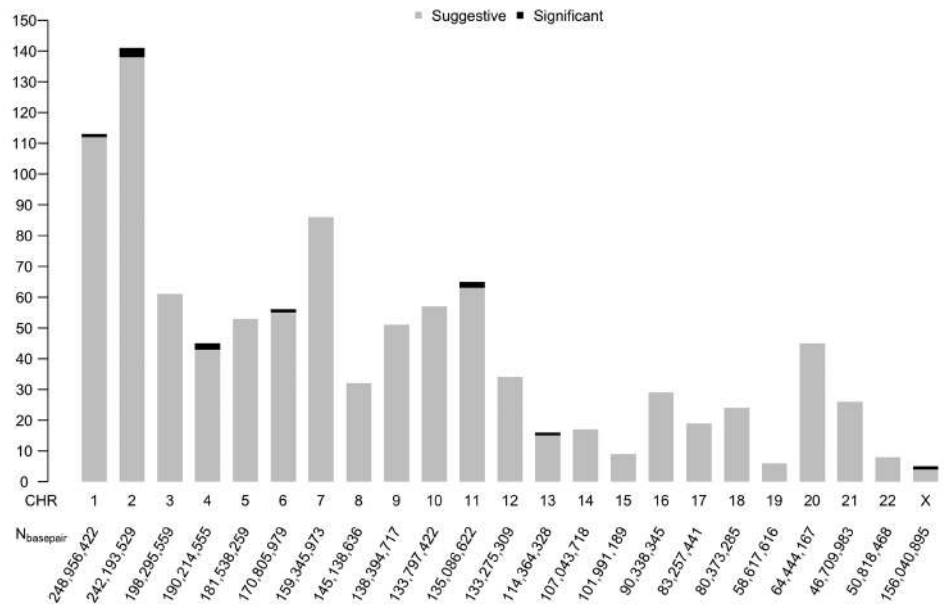


Figure 3. Number of suggestive and significant genetic variants associated with aggression-related traits at $P \leq 1x10^{-05}$ on different chromosomes reported in overviewed GWASs

Chromosomes are numbered according to length (indicated by number of base pairs under the x-axis). The x-axis shows chromosome number and length (in base pairs). $N_{studies} = 18$, $N_{variants} = 998$. Updated by Odintsova et al, 2022 (in press).

lie inside or close to multiple genes, including *LRRC7* (leucine rich repeat containing 7), involved in neuronal excitability and used as postsynaptic marker of hippocampal glutamatergic synapse integrity, and *STIP1* (stress induced phosphoprotein 1), involved in astrocyte differentiation and highly expressed in the brain. A second GWAS by Mick and colleagues[86] observed a nominal association of proneness to anger with the gene, involved in calcium influx and release in the post-synaptic density, and in long-term potentiation (*FYN* (FYN proto-oncogene, Src family tyrosine kinase)). McGue et al.[87] reported four SNPs associated with behavioural disinhibition including symptoms of CD and aggression, one of which (rs1368882; $P=1.90x10^{-06}$) was located inside the *GLIS1* (GLIS family zinc finger 1) gene responsible for a transcription factor that is involved in regulating the expression of numerous genes.

Recently, two larger studies attempted to identify genes associated with aggression or ASB by increasing power through the inclusion of multiple cohorts. Pappa and colleagues[88] collected a sample of 18988 children of 3 – 15 years old for meta-analysis and reported a near genome-wide significant locus on chromosome 2p12 ($P=5.3x10^{-08}$). This locus is in close proximity to two genes: *LRRTM4* (leucine-rich repeat transmembrane neuronal 4), which regulates excitatory synapse development, and *SNAR-H* (small NF90 (ILF3) associated RNA H), which is implicated in the transcription process and is expressed in neurons. They found 19 genes nominally related to aggression from gene-based tests,

phenotypes such as human height, but variance explained for psychiatric phenotypes is smaller. PGSs, based on the Ip et al.[89] GWAMA, explained between 0.04% and 0.44% of the variance in aggression of 7-year olds. Applications of PGS research on aggression are relatively rare, but can be a powerful method to investigate specific hypotheses on genetic influences. For example, van der Laan et al.[90], demonstrated that PGSs based on the discovery GWAS of aggression in children predict aggression across the life course, indicating long-lasting genetic influences and continuity of genetic effects. Barnes et al.[91] used PGSs to demonstrate gene-environment interaction; a PGS for aggressive behaviour was positively associated with incarceration risk, but only for participants who were born into a home where neither parent had completed high school.

Epigenome-wide association studies *

EWASs have become feasible through DNA methylation microarray detecting differentially methylated loci at cytosine-phosphate-guanine (CpG) positions in the genome and applied to DNA extracted from different tissues. DNA methylation is one of the epigenetic mechanisms that mediates effects of genetic variants in regulatory regions on gene expression[47,92]. DNA methylation can reflect influences of environmental exposures and influence of DNA polymorphisms[47]. The question of how environmental influences might be translated into intracellular information and molecular memory is addressed through studying epigenetic patterns related to different conditions and traits[93]. Methylation profiles are influenced by early exposures, such as prenatal maternal smoking, and may be triggered by early life adversity[30,42,45]. Epigenetic changes can be both cause and consequence of aggressive behavior[94 95] and research designs such as longitudinal studies or studies in discordant twin and sibling pairs are needed to shed light on causality. Unlike DNA markers, methylation profiles are tissue-specific. Some studies suggest that DNA methylation in some peripheric tissues may be useful to investigate brain-based phenotypes such as aggression, as it is associated with central nervous system methylation, e.g., in DNA extracted from peripheral blood[95] and from buccal cells[96].

Epigenetic alterations in genes involved in neural function have been observed in psychiatric disorders which incorporate aggression as a diagnostic criterion, such as childhood disruptive behavior disorders[97] and ASPD[98]. Small candidate gene studies and EWAS (sample size <25) have provided some evidence that DNA methylation differences in blood cells are associated with chronical physical aggression in children and adults[99–102]. The larger EWAS performed with microarray technologies on Illumina 450k and Illumina EPIC arrays found up to 16 more robust association at multiple testing correction (see **Table 4**). These studies were performed in cord blood, blood cells, and saliva cells, and included engagement in physical fights[94], intermittent explosive disorder measured by DSM-V and aggression scales[103], childhood aggression measured with Child Behavior Check List (CBCL)[104], and adult aggression measured with ASEBA self-report[105]. The functional analysis of the loci gives reason to suppose involvement of associated genes in neuronal development, immune and endocrine systems, and metabolic processes.

* Added in updated version Odintsova et al., 2022 (in press)

Van Dongen and colleagues[105] analyzed data from 20 discordant monozygotic twin pairs highly discordant for aggression measured by ASEBA, thereby controlling for age, sex, birth year, genetic confounding, and confounding by perinatal and early life factors (e.g., maternal genotype, parental age, maternal lifestyle during pregnancy, breastfeeding, same household, and nutrition). Aggression discordance was accompanied by small DNA methylation differences in blood at loci that could be related to etiology of aggressive behavior. CpGs showing a large difference in methylation were generally twin pair-specific. Methylation sites with a lower *P*-value were located more often than expected by chance in or near genes related to various central nervous system processes.

The largest epigenetic meta-analysis to date combined genome wide methylation data from 18 cohorts from different countries with data on aggressive behavior and closely related constructs in 15,324 participants[106]. Sixteen CpGs showed a robust association with aggressive behavior at stringent threshold for multiple testing correction at $P<1.2x10^{-7}$ across the peripheral blood meta-analysis and combined meta-analysis of DNA methylation profiles in peripheral blood and cord blood. Three genes whose expression levels were associated with top findings were previously linked to psychiatric or behavioral traits in GWAS: *FLOT1* (flottilin 1) (schizophrenia[107]), *TUBB* (tubulin beta class I) (schizophrenia[107]), and *PLXNA2* (plexin A2) (general risk tolerance[34]). Several other loci have functions in the brain, and six CpGs showed correlated methylation levels between blood and brain. Interestingly, CpGs associated with aggressive behavior at $P<1x10^{-5}$ have been associated with chemical exposures, smoking, cognition, metabolic traits, and genetic variation. This suggests a role of the environment on molecular variation associated with behavioral traits. However, causal relationships, the mediation role of DNA methylation in gene x environment interactions in trait formation, and confounding effects, as well as more robust DNA methylation signatures, remain to be unraveled.

Studies of human aggression in other omics domains *

Compared with (epi-)genetic studies, studies investigating the association of other omics domains with aggressive behavior remain relatively rare. Here we give several examples, describing a metabolomics studies for aggressive behavior[108,109], a study that combines a transcriptome-wide association study (TWAS) with proteomics for violence[110], and a microbiome study for violent behavior in schizophrenia[111].

In a sample of 725 young adults, the association of 11 plasma metabolites with aggressive behavior was tested and a negative association between 3-hydroxybutyrate and aggression was reported, that was consistent across ages and raters, with both teacher (age 12) and self (age 14) aggression ratings contributing uniquely to this association[109]. Replication of this finding in 960 adults showed the same trend of lower 3-hydroxybutyrate levels being associated with higher aggression levels, but was not significant. The concentrations of urinary metabolites and a number of classic biomarkers were assessed in children scoring low and high on aggressive behavior[108]. In the discovery sample of 785 children, concentrations of O-phosphoserine and gamma-L-glutamyl-L-alanine had a significantly positive association with aggression scores. In a replication in 189 twins

* Added in updated version Odintsova et al., 2022 (in press)

discordant for aggression, no significant differences between the low and high scoring twins were seen. Validation in 183 clinical cases and 184 matched controls of the top 5 urinary biomarkers (gamma-glutamylglutamine, L-arginine, glyceric acid, creatinine, and succinic acid) with congruent directions of effect in the discovery and replication analyses found no significant associations of these biomarkers with childhood aggression.

A TWAS in cortical neurons and astrocytes, derived from induced pluripotent stem cells (iPSC), of six incarcerated extremely antisocial and violent offenders, three individuals with substance abuse, and six healthy controls, described 39 transcripts that were differentially expressed in violent offenders[110]. In neurons, 2 genes (*RPL10P9* (ribosomal protein L10 pseudogene 9) and *ZNF132* (zinc finger protein 132)) were upregulated in violent offenders when compared to healthy controls. When comparing the violent offenders to healthy controls and individuals with substance abuse a further 8 upregulated and 19 downregulated genes were identified. In astrocytes, 2 genes, *RPL10P9* and *MT-RNR2* (mitochondrially encoded 16S RNA), were upregulated in cases compared to healthy controls, and a further 7 upregulated and 8 downregulated genes were found when comparing violent offenders to all others. The strongest expression differences were observed for the *RPL10P9, ZNF132, CDH5* (cadherin 5), and *OPRD1* (opioid receptor delta 1) genes, and these genes explained 30%–92% of the variance in psychopathic symptoms. This study also investigated the proteome and phosphoproteome, i.e., those proteins that have received a phosphate group during posttranslational modification[112], in the iPSC-derived cortical neurons. These analyses identified 3 proteins and 7 phosphoproteins that were upregulated in cases, and another 3 proteins and 7 phosphoproteins that were downregulated. The largest expression differences were seen for *OPCML* (opioid-binding protein/cell-adhesion molecule), *PSMD3* (proteasome 26S subunit non-ATPase 3), *PEG10* (paternally expressed 10), and *PCDH19* (protocadherin 19).

When comparing schizophrenia patients with violent behavior (N = 53) to patients without violent behavior (N = 23), univariate analyses of plasma metabolites identified 3 metabolites that were significantly increased, and 16 metabolites that were significantly decreased[136]. With these 19 metabolites two multivariate classifiers were build that selected 10 and 5 metabolites as biomarkers for schizophrenia with violent behavior. A study investigating differences in the gut microbial structure in schizophrenia patients with (N = 26) and without violent behavior (N = 16), detected 59 differential microbial taxonomic compositions between the groups using 16S rRNA sequencing[111]. Across the class, order, family, and genus level many microbiomes were observed in patients without violent behavior but not in patients with violent behavior.

Discussion

Aggression has a considerable genetic component, as indicated by decades of behaviour genetics research. However, no genomic variants have (yet) been identified. In our review covering GWASs on human aggression, only 4 out of 17 studies reported genome-wide significant hits in primary or replication samples[27,72–74]. In the reviews on aggression and GWASs, several explanations are offered for the discrepancy between heritability

TABLE 4.
Overview of epigenome-wide suggestive and significant associations with aggression-related traits at $P \leq 1x10^{-05}$ per EWAS performed with Illumina 450k and EPIC

Study	Phenotype	Sample	Tissue and Platform of DNA methylation	N_{CpGs}	Genes	Summary of main findings
Van Dongen et al., 2015	Aggressive behavior, ASEBA Adult Self-Report	*N*=2029. Males 31%. Mean age at aggression measurement: 38 years. Mean age at blood sampling: 36.4. *N*=40 (N_{cases}=20, $N_{controls}$=20). Monozygotic twins discordant for aggression. Males 15%. Mean age at aggression measurement 36.7 years. Mean age at blood draw 34.5. The Netherlands Twin Register	PBC, Illumina 450k PBC, Illumina 450k	8 at $P<5x10^{-05}$ 3 at $P<5x10^{-05}$	*TRPS1, PARD6G-AS1, RNU5F-1, SEC23A, C9orf5, C21orf13, OSR1, POPDC3 RAB39, SIGLEC10, PREP*	No associations at genome-wide significant threshold or FDR 5%. The two highest ranking CpGs located on chromosomes 8 and 18. No associations at genome-wide significant threshold or FDR 5%. Differentially methylated CpGs with methylation difference >30% 24 CpGs, more than 15% 291 CpGs. Top sites were close to *RAB39, SIGLEC10*, and *PREP*
Montalvo-Ortiz et al., 2018a		*N*=44 (N_{cases}=22, $N_{controls}$=22). Males 52%. Mean age 37.	PBC, Illumina 450k	27 at $P<5x10^{-05}$	*PTPRN2, HK1, TYMS, CYFIP1, PLD2, ALDOA, YME1L1, FEZF2, KDELC2, SAR1A, MAD1L1, DIABLO, C6orf136, SORD, EOGT, PRDM2, NR1H2, GBP4, TMEM180, ADAMTS17, TMEM81, LOC100505921, CPLX2, FN3KRP, BMPR1A, INHBA*	No associations at genome-wide significant threshold. Differentially methylated CpGs at P<5x10-05 were associated with genes involved in inflammatory/immune system, endocrine system, and neuronal differentiation

Table 4. (continued)

Study	Phenotype	Sample	Tissue and Platform of DNA methylation	N_{CpGs}	Genes	Summary of main findings
Cecil et al., 2018	Physical aggression, engagement in physical fights, Youth Risk Behavior Survey ("During the past 12 months, how many times were you in a physical fight?")	*N*=119 (N_0=88, $N_{1\ time}$ =16, $N_{2\text{-}5\ times}$ = 11, $N_{>5times}$ = 4). High risk group. Males 47%. Age range: 16-24 years. Replication: adults based on DNA methylation in blood T-cells	Buccal swabs, Illumina 450k	4 at $P<5x10^{-07}$	*TRH, GAK, RIC3, CYFIP2 (top CpGs), DRD4* (differentially methylated region)	Enrichment of top CpGs for a range of biological processes, including axon guidance, hormone metabolic processes, behavioral regulation, dopamine receptor signaling, cytokine secretion, and drug response. One DMR in DRD4 gene was identified and showed strong cross-tissue concordance with both blood and brain. CpGs within this region were also differentially methylated in an independent sample of adults with DNA methylation in blood
Montalvo-Ortiz et al., 2019	CBCL	*N*=105 (N_{cases} 39%). Males 44%. Age range: 9-15 years	Buccal swabs, Illumina 450k	8 at $P<5x10^{-07}$	*HPCAL4*	From genome-wide significant sites (at $P<5.0x10^{-7}$), cg18438793 located within *HPCAL4* gene is of importance due to its role in neuronal development and calcium signaling and brain-specific gene expression. Other sites (P value range: $1x10^{-7} – 1x10^{-10}$) are involved in axon guidance, myelination and transcriptional regulation
van Dongen et al., 2021	CBCL, SDQ, MNPI aggression scale, ASEBA adult self-report, DSM-IV Conduct Disorder Symptom	*N*=15,324. 21 cohorts. Males 45%. Age range PBC: 7-68 years. Age range cord blood: 4-7 years	PBC + cord blood, Illumina 450k, EPIC	13 at $P<1.2x10^{-07}$	*AHRR, GFI1, GFI11, MYO1G,*	Thirteen CpGs were significantly associated with aggression in peripheral blood and 13 in the combined meta-analysis of blood and cord blood (16 unique sites). 83% of the FDR 5% CpGs from peripheral blood EWAS showed the same direction of association with childhood aggression, but no epigenome-wide significant sites were found. No CpGs were epigenome-wide

Table 4. (continued)

Study	Phenotype	Sample	Tissue and Platform of DNA methylation	N_{CpGs}	Genes	Summary of main findings
	Scale, MPQ aggression scale, Hunter-Wolf aggressive behavior scale					significant in the adjusted meta-analysis. Top-CpGs have been associated with chemical exposures, smoking, cognition, metabolic traits, and genetic variation. At six CpGs, DNA methylation variation in blood mirrors variation in the brain. DRD4 region (Cecil et al, 2018) was not replicated. DNA methylation scores explained 0.29% of the variance in aggression.

P, statistical significance assessed considering for the number of tested CpGs; ASEBA, Achenbach System of Empirically Based Assessment scales; BPAQ, Buss-Perry Aggression Questionnaire; CBCL, Child Behavioral Checklist; CPA, chronical physical aggression; DSM, Diagnostic and Statistical Manual of Mental Disorders; MNPI, Multidimensional Peer Nomination Inventory aggression scale; MPQ, Multidimensional Personality Questionnaire; PBC, Peripheral blood cells; SBQ, Social Behavior Questionnaire; SDQ, Strengths and Difficulties Questionnaire.
Note: Added in updated version Odintsova et al., 2022 (in press)

estimates in behavioural and molecular genetic studies; for example, the heterogeneous, context-dependent, and developmental nature of aggression, but foremost, small sample sizes. Fortunately, these limitations can be remedied, and provide future directions for research.

Heritability estimates are 50% for aggression (by Miles and Carey[113]), and 41% for ASB (by Rhee and Waldman[114]), and are confirmed in multiple empirical studies. Moderation, or any genotype x environment effects seem small, and most pronounced for nonaggressive ASB[115]. Early linkage studies on aggression indicated as potential loci chromosomes 1 by Criado et al.[64], 2, and 19 by Dick et al.[63]. GWAS findings in our review confirm loci on chromosomes 1 and 2 which gave more associated variants and significant results. The X- and Y-chromosomes did not give evident results, but note that these are often excluded from GWAS projects[27].

How to address nonsignificant findings in GWAS studies on psychiatric problems is a pressing issue. Opinions are divided on what approach is most optimal to define phenotypes for GWAS analyses. Some believe that reduction of phenotypic heterogeneity could lead to more genome-wide significant findings[8,40,116]. This view is supported by the GWASs covered in this review that did find genome-wide significant hits. These relatively underpowered studies (Nrange from 2185 to 6220 participants) focus on individuals with severe antisocial behaviour and specific types of aggression: individuals with DSM-defined CD symptoms[72], cannabis-induced physical aggression[74], and criminal offenders with antisocial personality disorder[73]. Two studies were conducted in specific samples; exclusively male, with associations only in African-American subgroup[74], and predominantly male (89% of cases) and ethnically homogeneous[73].

In contrast, other researchers propose a broader approach which includes more lenient phenotypes[117,118]. This lenient phenotyping approach has already achieved success in depression research, for example, although here the value of minimal versus broader phenotyping is debated as well[37]. The three largest GWASs on aggression that were covered by this review used broad measures of childhood aggression[88,89] and ASB[27]. Pappa and colleagues (2016) found no significant hits, but several promising loci on chromosomes 2, 3, 6 and 17 (minimum P=5.3x10^{-08}). Ip and colleagues[89] reported 185 suggestive SNPs and three significant gene-associations in gene-based tests. Tielbeek and colleagues (2017) reported three significant hits for the sex-stratified GWAMAs.

To identify 80% of all causal SNPs, depending on the extent of SNP heritability, between 10^5 and 10^7 (100 000 – 10 000 000) independent subjects would be required[119]. This means that, with sample sizes 10 time less than the lower bound, current GWASs were clearly underpowered. At present, several initiatives are under way to collaborate in achieving larger sample sizes. One example of a large collaborative project is the ACTION consortium (Aggression in Children: Unraveling gene-environment interplay to inform Treatment and InterventiON strategies: http://www.action-euproject.eu/), which has brought together over 30 cohorts with childhood data on aggression for GWAS, EWAS and biomarker studies.

Concerns regarding heterogeneity and the impact of different phenotype definitions are empirical questions, which are currently also being asked in other GWAS of psychiatric

which include *LRRTM4, PDSS2* (decaprenyl diphosphate synthase subunit 2), *TRIM27* (tripartite motif containing 27), *MRC1* (mannose receptor C-type 1), *MECOM* (MDS1 and EVI1 complex locus), and *CASC17* (cancer susceptibility 17).

Another larger study by Tielbeek and colleagues[27] focused on the broader ASB phenotype in 16400 individuals. The overall GWAMA found no hits, but sex-stratified GWAMAs returned three genome-wide significantly associated SNPs (minimum P=1.95x10^{-08}), but failed to identify significant genes. This suggested that there might be sex-specific genetic effects on ASB and focusing on a more specific phenotype could improve chances of findings significant results.

Thus, nominal genome-wide associations (P<1x10^{-05}) have been found in genes involved in a wide variety of biological systems: the immune system, the endocrine system, pathways involved in neuronal development and differentiation and synaptic plasticity. These findings have not been replicated across GWASs, but some studies reported the same genes independently: *NTM*[81,83] and *RBFOX1(A2BP1)*[76,77].

Finally, Ip and colleagues[89] conducted the largest GWAMA of aggression to date. A total of 328,935 observations from 87,485 individuals were included in the analyses. Observations came from multiple raters (teachers, parents, and self-reports). No single SNP-based genome-wide significant hits were observed. Three significant genes were observed in gene-based analysis: *ST3GAL3* (ST3 beta-galactoside alpha-2,3-sialyltransferase 3) (P=1.6x10^{-06}), *PCDH7* (protocadherin 7) (P=2.0x10^{-06}), and *IPO13* (importin 13) (P=2.5x10^{-06}). Lead SNPs in all three implicated genes have previously been associated with educational traits.

In summary, GWASs show that genome-wide significant and/or suggestive associations between aggression-related traits and SNPs are found on all chromosomes. As shown in **Figure 3**, over 75% of suggestive associations were found on chromosomes 1 to 11. In other words, most suggestive associations are found on the largest chromosomes, which illustrates the polygenic aetiology of aggression. The majority of suggestive SNPs on chromosome 7 were reported in the sample of African ancestry[74]. The genome-wide significant associations are located on chromosomes 1, 2, 4, 6, 11, 13, and X. Nominal genome-wide associations (P<1x10^{-05}) have been found in genes involved in a wide variety of biological systems: the immune system, the endocrine system, pathways involved in neuronal development and differentiation, and synaptic plasticity. These findings have not been replicated across GWAS, but some studies reported the same genes independently: NTM (neurotrimin)[81,83], and *RBFOX1(A2BP1)* (RNA binding fox-1 homolog 1)[76,77].

By combining the effects of multiple SNPs into a score, a larger proportion of variance in aggression is likely be explained compared to the variance that is captured by a SNP alone. Polygenic scores (PGSs) translate GWAS results back to the individual level, by summing the association effects of all effect alleles that are present in an individual. Thus, a PGS is a measure of genetic propensity for a certain outcome at the individual level. PGSs can contribute to the prediction of aggression, but depend on the effect sizes of the associations in the discovery GWAS. Currently such predictions work reasonably well for

disorders[37]. Questions regarding whether different phenotypes reflect the same underlying genetic liability can be resolved, once well-powered GWAS are available, by estimation of genetic correlations among different phenotype definitions of aggression and can also be addressed through genetic modelling of twin and family data. For example, Hendriks et al.[38] analyzed twin data collected by multiple instruments, commonly employed to measure aggression in children. While phenotypic correlations between different aggression scales could be low, a genetic multivariate analysis of these data showed high genetic correlations among different instruments. Such observations mean that different instrument tap into the same genetic liability and therefore are appropriate for simultaneous analysis in GWAS.

A limitation of human epigenetic and transcriptome studies concerns the use of peripheral tissues to study genetic expression for a brain-based phenotype. Here, animal studies may provide additional insights. A study of aggression in silver fox strains selectively bred over many generations (in experiment lasting 60 years) for increased aggression and sociableness, and reduced fear to humans, revealed unique gene and transcript profiles of aggression[120]. Behavioral differences between tame and aggressive foxes altered the anatomy of distributed gray matter networks, which included the hypothalamus[121]. Around 70 differentially expressed genes were identified in the hypothalamus, a key brain region involved in regulation of aggression in animals[122] and humans[123]. Functional analyses of genes altered between tamed and aggressive foxes highlighted underlying biological processes that include neuronal development, differentiation, migration, and synaptic plasticity, immunity, lipid and carbohydrate metabolism, extracellular matrix organization, cell interaction, and several signaling pathways[124].

We should recognize that the nature-nurture debate has moved on from the question whether aggressive behavior is heritable to the discovery of the biological bases of aggression. This is currently attempted by investigating the association of aggression with DNA variation and relevant biological pathways. It is expected that GWAS with larger or combined datasets will improve our understanding of the mechanisms of gene regulation of aggression. Future work should determine if genes mediating aggression pathways are enriched in the polygenic background of disorders associated with aggression.

The majority of the studies described in this chapter focused on a single omics layer. By considering each of these omics' layers separately, these studies identified genes, CpGs, proteins, metabolites, and microbes that associate with aggression or related traits. These studies have provided valuable insights into the underlying biology of aggression and related traits, but single omics studies do not take into account the complex interplay among the various omics layers[125]. Combining multiple layers of omics data (e.g., transcripts, CpGs, or metabolites) in a multi-omics study is a promising approach towards a more comprehensive biological understanding of aggression and related traits. Generally, two methods, simultaneous and step-wise integration, can be distinguished. Simultaneous, or parallel, integration methods combine multiple omics layers in a single analysis, accounting for correlations between the layers. Step-wise, or sequential, integration techniques first analyze the different omics layers separately and then combine and integrate the results[126].

Both methods come with their own advantages and limitations[126]. One limitation of simultaneous integration methods is the need for omics data from the same individuals, which is often not available. In contrast, sequential integration methods can integrate data from different sources and thus allows for the large-scale multi-omics analyses. As a consequence, many recent GWAS or EWAS include some sequential multi-omics analyses as follow-up analyses. For example, leveraging on Genotype-Tissue Expression GWAS findings can be annotated with additional information and thereby identify biologically relevant systems. Such biological annotation involving information from expression quantitative trait loci (eQTL) looks at SNPs that have been associated with gene expression levels. Once genome-wide hits are found for e.g., aggression, integration of these findings with known eQTLs could identify genes that are of biological interest[128,129].

Future directions

We should recognize that the nature-nurture debate has moved on from the question whether aggressive behaviour is heritable to the discovery of the biological bases of aggression. This is currently achieved by investigating aggression's relation to genes, SNPs, and relevant biological pathways. It is expected that GWASs with larger or combined datasets will improve our understanding of the mechanisms of gene regulation of aggression. Individual GWASs on aggression and aggression-like traits are still limited in terms of explaining variation in the population, but ongoing GWASs and other efforts, e.g. in epigenetics and biomarker studies are likely provide insight into the aetiology of aggressive behaviour. Expansion of disease gene maps[130] by including aggression-related traits into, for example, OMIM datasets can help in future analyses of underlying cellular network-based relationships between genes and functional modules of aggressive behaviour, and future work should determine if genes mediating aggression pathways are enriched in the polygenic background of disorders associated with aggression.

Also, leveraging on Genotype-Tissue Expression (GTEx[127]) GWAS findings can be annotated with additional information and thereby identify biologically relevant systems. One particularly interesting source of biological annotation revolves expression quantitative trait loci (eQTL), i.e. SNPs that have been associated with gene expression levels. Once genome-wide hits are found, overlapping these with known eQTLs could identify genes that are of biological interest[128,129,131].

Risk-factors-based prediction of aggressive behavior and classification accuracy is expected to be improved with (epi)genetic and other omics biomarkers, i.e., characteristics that are "objectively measured and evaluated as an indicator of normal biological processes, pathological processes or biological responses to a therapeutic intervention"[132]. Non-causal biomarkers can be used in diagnostics, while causal biomarkers are potential target for prevention and treatment. However, more work is needed to find links of gene expression between brain and other tissues, to detect the causality between the signatures and phenotype, and to develop reliable biomarkers in easily accessible peripheral tissues.

Systematic reviews with automated functions

The workload on selection process of researchers in our systematic review was around 60 hours (screening and selecting relevant papers from list of 2 069 records). By using automated procedures to screen for relevant literature for inclusion in systematic reviews, it was possible to save 39.1% (23.5 hours) of reading/scanning time. The downside of automated methods is that relevant literature can be missed. On the other hand, even an expert reviewer might omit studies that the automated procedures include. Optimization of the expert reviewer is covered by education and training, whereas optimization of automated selection is under active development[133–135]. We opted for a recent approach that utilizes a machine learning algorithm to obtain a selection of papers that could be relevant for this systematic review.

Although the automated systematic review tool we applied is quite new and is still under active development, we found that applying the machine learning approach as implemented in the software "Automated systematic reviews by using Deep Learning and Active Learning"[28] could be indeed of considerable aid to the researcher performing a systematic review solving problems of missed literature in screening phase due to human errors or excluded by searching algorithms.

For the benefit of further developments in automated selection approaches aiding the review process, we advise review authors to supply their search results as additional information to their work. These results can then serve for further refinement of literature search models. This would avoid double work across research groups, create a comprehensive overview of aggression literature, and increase our understanding of the genetic nature of human aggression.

Conclusions

In summary, aggression in humans is a heritable trait, whose genetic basis largely remains to be uncovered. With increases in sample size, we expect aggression to behave like other complex human traits for which GWAS have been successful. There are several efforts to achieve powerful GWAS: merging samples in consortia, replication strategies, searching for close phenotypes from other domains associated with aggression for sample extension, developing new approaches of partitioning genetic heterogeneity and sample stratification.

References

*Articles included in systematic review are indicated: * reviews (Table 2), ** genome-wide association studies (Table 3), and *** epigenome-wide association studies (Table 4).*

1. Tuvblad, C. & Baker, L. A. Human aggression across the lifespan: genetic propensities and environmental moderators. *Adv Genet* **75**, 171–214 (2011). *
2. Veroude, K. *et al.* Genetics of aggressive behavior: An overview. *Am J Med Genet B Neuropsychiatr Genet* **171B**, 3–43 (2016). *
3. Anderson, C. A. & Bushman, B. J. Human Aggression. *Annual Review of Psychology* **53**, 27–51 (2002).
4. Zhang-James, Y. & Faraone, S. v. Genetic architecture for human aggression: A study of gene-phenotype relationship in OMIM. Am J Med Genet B *Neuropsychiatr Genet* **171**, 641–649 (2016). *
5. Boomsma, D. I. *et al.* Aggression in children: unravelling the interplay of genes and environment through (epi)genetics and metabolomics. *Journal of Pediatric and Neonatal Individualized Medicine* **4**, (2015).
6. Lindenfors, P. & Tullberg, B. S. Evolutionary aspects of aggression the importance of sexual selection. *Adv Genet* **75**, 7–22 (2011).
7. Hawley, P. H., Little, T. D. & Rodkin, P. C. *Aggression and adaptation: the bright side to bad behavior.* (L. Erlbaum Associates, 2007).
8. Anholt, R. R. & Mackay, T. F. Genetics of aggression. *Annu Rev Genet* **46**, 145–164 (2012). *
9. Bukowski, W. M., Laursen, B. P. & Rubin, K. H. *Handbook of peer interactions, relationships, and groups.* (2011).
10. Dodge, K. A., Lochman, J. E., Harnish, J. D., Bates, J. E. & Pettit, G. S. Reactive and proactive aggression in school children and psychiatrically impaired chronically assaultive youth. *Journal of Abnormal Psychology* **106**, 37–51 (1997).
11. Hawley, P. H. Strategies of control, aggression, and morality in preschoolers: An evolutionary perspective. *Journal of Experimental Child Psychology* **85**, 213–235 (2003).
12. Voulgaridou, I. & Kokkinos, C. M. Relational aggression in adolescents: A review of theoretical and empirical research. *Aggression and Violent Behavior* **23**, 87–97 (2015).
13. Wright, J. C., Giammarino, M. & Parad, H. W. Social status in small groups: Individual–group similarity and the social "misfit." *Journal of Personality and Social Psychology* **50**, 523–536 (1986).
14. Penn, J. K. M., Zito, M. F. & Kravitz, E. A. A single social defeat reduces aggression in a highly aggressive strain of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 12682–6 (2010).
15. Polman, H., Orobio de Castro, B., Koops, W., van Boxtel, H. W. & Merk, W. W. A Meta-Analysis of the Distinction between Reactive and Proactive Aggression in Children and Adolescents. *Journal of Abnormal Child Psychology* **35**, 522–535 (2007).
16. Brunner, H. G., Nelen, M., Breakefield, X. O., Ropers, H. H. & van Oost, B. A. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. Science (New York, N.Y.) 262, 578–80 (1993).
17. Caspi, A. *et al.* Role of genotype in the cycle of violence in maltreated children. *Science (New York, N.Y.)* **297**, 851–4 (2002).
18. Haberstick, B. C. *et al.* Monoamine oxidase A (MAOA) and antisocial behaviors in the presence of childhood and adolescent maltreatment. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **135B**, 59–64 (2005).
19. Young, S. E. *et al.* Interaction Between MAO-A Genotype and Maltreatment in the Risk for Conduct Disorder: Failure to Confirm in Adolescent Patients. *American Journal of Psychiatry* **163**, 1019–1025 (2006).
20. Foley, D. L. *et al.* Childhood Adversity, Monoamine Oxidase A Genotype, and Risk for ConductDisorder. *Archives of General Psychiatry* **61**, 738 (2004).

21. Nilsson, K. W., Aslund, C., Comasco, E. & Oreland, L. Gene-environment interaction of monoamine oxidase A in relation to antisocial behaviour: current and future directions. *J Neural Transm (Vienna)* **125**, 1601–1626 (2018).
22. Kim-Cohen, J. *et al.* MAOA, maltreatment, and gene–environment interaction predicting children's mental health: new evidence and a meta-analysis. *Molecular Psychiatry* **11**, 903–913 (2006).
23. Dudbridge, F. Polygenic Epidemiology. *Genetic Epidemiology* **40**, 268–272 (2016).
24. Hagenbeek, F. A. et al. Discovery of biochemical biomarkers for aggression: A role for metabolomics in psychiatry. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **171**, 719–732 (2016).
25. Wild, C. P. Complementing the Genome with an "Exposome": The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiology Biomarkers & Prevention* **14**, 1847–1850 (2005).
26. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics* **101**, 5–22 (2017).
27. Tielbeek, J. J. *et al.* Genome-Wide Association Studies of a Broad Spectrum of Antisocial Behavior. *JAMA Psychiatry* **74**, 1242–1250 (2017). **
28. Automated systematic reviews by using Deep Learning and Active Learning. <*https://github.com/msdslab/automated-systematic-review*> (2019).
29. Fernandez-Castillo, N. & Cormand, B. Aggressive behavior in humans: Genes and pathways identified through association studies. *Am J Med Genet B Neuropsychiatr Genet* **171**, 676–696 (2016). *
30. Tremblay, R. E., Vitaro, F. & Cote, S. M. Developmental Origins of Chronic Physical Aggression: A Bio-Psycho-Social Model for the Next Generation of Preventive Interventions. *Annu Rev Psychol* **69**, 383–407 (2018). *
31. Raine, A. A neurodevelopmental perspective on male violence. *Infant Ment Health J* **40**, 84–97 (2019).
32. Baud, P. Personality traits as intermediary phenotypes in suicidal behavior: Genetic issues. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* **133C**, 34–42 (2005). *
33. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* (American Psychiatric Association, 2013). doi:10.1176/appi.books.9780890425596.
34. Karlsson Linnér, R. *et al.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nature genetics* **51**, 245–257 (2019).
35. Barr, P. B. & Dick, D. M. The Genetics of Externalizing Problems. *Current topics in behavioral neurosciences* **47**, 93–112 (2020).
36. Ferguson, C. J. Genetic contributions to antisocial personality and behavior: a meta-analytic review from an evolutionary perspective. *J Soc Psychol* **150**, 160–180 (2010).
37. Cai, N. *et al.* Minimal phenotyping yields GWAS hits of low specificity for major depression. *bioRxiv* 440735 (2019) doi:10.1101/440735.
38. Hendriks, A. M. *et al.* Content, diagnostic, correlational, and genetic similarities between common measures of childhood aggressive behaviors and related psychiatric traits. *Journal of child psychology and psychiatry, and allied disciplines* **61**, 1328–1338 (2020).
39. Waltes, R., Chiocchetti, A. G. & Freitag, C. M. The neurobiological basis of human aggression: A review on genetic and epigenetic mechanisms. *Am J Med Genet B Neuropsychiatr Genet* **171**, 650–675 (2016). *
40. Runions, K. C. *et al.* Serotonin and aggressive behaviour in children and adolescents: a systematic review. *Acta Psychiatr Scand* **139**, 117–144 (2019).

41. Moffitt, T. E. Genetic and environmental influences on antisocial behaviors: evidence from behavioral-genetic research. *Adv Genet* **55**, 41–104 (2005). *
42. Provencal, N., Booij, L. & Tremblay, R. E. The developmental origins of chronic physical aggression: biological pathways triggered by early life adversity. *J Exp Biol* **218**, 123–133 (2015). *
43. Davydova, J. D., Litvinov, S. S., Enikeeva, R. F., Malykh, S. B. & Khusnutdinova, E. K. Recent advances in genetics of aggressive behavior. *Vavilov Journal of Genetics and Breeding* **22**, 716–725 (2018). *
44. Tremblay, R. E. Developmental origins of disruptive behaviour problems: the "original sin" hypothesis, epigenetics and their consequences for prevention. *J Child Psychol Psychiatry* **51**, 341–367 (2010).
45. Manchia, M. & Fanos, V. Targeting aggression in severe mental illness: The predictive role of genetic, epigenetic, and metabolomic markers. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* **77**, 32–41 (2017). *
46. Curry, A. A painful legacy. *Science* **365**, 212–215 (2019).
47. Van Dongen, J. *et al.* Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nature Communications* **7**, 11115 (2016).
48. Achenbach, T. M., Ivanova, M. Y. & Rescorla, L. A. Empirically based assessment and taxonomy of psychopathology for ages 1½–90+ years: Developmental, multi-informant, and multicultural findings. *Comprehensive Psychiatry* **79**, 4–18 (2017).
49. Goodman, A., Lamping, D. L. & Ploubidis, G. B. When to Use Broader Internalising and Externalising Subscales Instead of the Hypothesised Five Subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British Parents, Teachers and Children. *Journal of Abnormal Child Psychology* **38**, 1179–1191 (2010).
50. Buss, A. H. & Durkee, A. An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology* **21**, 343–349 (1957).
51. Van der Sluis, S., Posthuma, D., Nivard, M. G., Verhage, M. & Dolan, C. v. Power in GWAS: lifting the curse of the clinical cut-off. *Molecular psychiatry* **18**, 2–3 (2013).
52. Loehlin, J. C. History of Behavior Genetics. in *Handbook of Behavior Genetics* 3–11 (Springer New York, 2009). doi:10.1007/978-0-387-76727-7_1.
53. Trivedi, D. K., Hollywood, K. A. & Goodacre, R. Metabolomics for the masses: The future of metabolomics in a personalized world. *New Horizons in Translational Medicine* **3**, 294–305 (2017).
54. Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. *Nature Reviews Genetics* **3**, 872–882 (2002).
55. Craig, I. W. & Halton, K. E. Genetics of human aggressive behaviour. *Hum Genet* **126**, 101–113 (2009).
56. Rhee, S. H. & Waldman, I. D. Genetic and environmental influences on aggression. in *Human aggression and violence: Causes, manifestations, and consequences* (eds. Shaver, P. R. & Mikulincer, M.) 143–163 (American Psychological Association, 2011).
57. Gard, A. M., Dotterer, H. L. & Hyde, L. W. Genetic influences on antisocial behavior: recent advances and future directions. *Curr Opin Psychol* **27**, 46–55 (2018). *
58. Porsch, R. M. *et al.* Longitudinal heritability of childhood aggression. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **171**, 697–707 (2016).
59. Burt, S. A., Klump, K. L., Gorman-Smith, D. & Neiderhiser, J. M. Neighborhood Disadvantage Alters the Origins of Children's Nonaggressive Conduct Problems. *Clinical Psychological Science* **4**, 511–526 (2016).
60. Boutwell, B. B., Beaver, K. M., Barnes, J. C. & Vaske, J. The developmental origins of externalizing behavioral problems: Parental disengagement and the role of gene–environment interplay. *Psychiatry Research* **197**, 337–344 (2012).

61. Hicks, B. M., South, S. C., DiRago, A. C., Iacono, W. G. & McGue, M. Environmental Adversity and Increasing Genetic Risk for Externalizing Disorders. *Archives of General Psychiatry* **66**, 640 (2009).
62. Vink, J. M. *et al.* Sex Differences in Genetic Architecture of Complex Phenotypes? *PLoS ONE* **7**, e47371 (2012).
63. Dick, D. M. *et al.* A genome-wide screen for genes influencing conduct disorder. *Molecular Psychiatry* **9**, 81–86 (2004).
64. Criado, J. R., Gizer, I. R., Slutske, W. S., Phillips, E. & Ehlers, C. L. Event-related oscillations to affective stimuli: Heritability, linkage and relationship to externalizing disorders. *Journal of Psychiatric Research* **46**, 256–263 (2012).
65. Salvatore, J. E. & Dick, D. M. Genetic influences on conduct disorder. *Neurosci Biobehav Rev* **91**, 91–101 (2018). *
66. Duncan, L. E. & Keller, M. C. A Critical Review of the First 10 Years of Candidate Gene-by-Environment Interaction Research in Psychiatry. *American Journal of Psychiatry* **168**, 1041–1049 (2011).
67. Luo, X. *et al.* Does refining the phenotype improve replication rates? A review and replication of candidate gene studies on Major Depressive Disorder and Chronic Major Depressive Disorder. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **171**, 215–236 (2016).
68. Bosker, F. J. *et al.* Poor replication of candidate genes for major depressive disorder using genome-wide association data. *Molecular Psychiatry* **16**, 516–532 (2011).
69. Sham, P. C. & Purcell, S. M. Statistical power and significance testing in large-scale genetic studies. *Nature Reviews Genetics* **15**, 335–346 (2014).
70. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291–295 (2015).
71. Beaver, K. M., Connolly, E. J., Nedelec, J. L. & Schwartz, J. A. On the Genetic and Genomic Basis of Aggression, Violence, and Antisocial Behavior. In: Oxford Handbooks Online (Ed. Hopcroft R.L.). (2018) *
72. Dick, D. M. *et al.* Genome-wide association study of conduct disorder symptomatology. *Molecular Psychiatry* **16**, 800–808 (2011). **
73. Rautiainen, M. R. *et al.* Genome-wide association study of antisocial personality disorder. *Transl Psychiatry* **6**, e883 (2016). **
74. Montalvo-Ortiz, J. L. *et al.* Translational studies support a role for serotonin 2B receptor (HTR2B) gene in aggression-related cannabis response. *Mol Psychiatry* **23**, 2277–2286 (2018). **
75. Kaye, S. *et al.* Upregulation of Early and Downregulation of Terminal Pathway Complement Genes in Subcutaneous Adipose Tissue and Adipocytes in Acquired Obesity. *Frontiers in Immunology* **8**, 545 (2017).
76. Anney, R. J. *et al.* Conduct disorder and ADHD: evaluation of conduct problems as a categorical and quantitative trait in the international multicentre ADHD genetics study. *Am J Med Genet B Neuropsychiatr Genet* **147B**, 1369–1378 (2008). **
77. Sonuga-Barke, E. J. *et al.* Does parental expressed emotion moderate genetic effects in ADHD? An exploration using a genome wide association scan. *Am J Med Genet B Neuropsychiatr Genet* **147B**, 1359–1368 (2008). **
78. Viding, E. *et al.* In search of genes associated with risk for psychopathic tendencies in children: a two-stage genome-wide association study of pooled DNA. *Journal of Child Psychology & Psychiatry* **51**, 780–788 (2010). **
79. Tielbeek, J. J. *et al.* Unraveling the genetic etiology of adult antisocial behavior: a genome-wide association study. *PLoS One* **7**, e45086 (2012). **
80. Salvatore, J. E. *et al.* Genome-wide association data suggest ABCB1 and immune-related gene sets may be involved in adult antisocial behavior. *Transl Psychiatry* **5**, e558 (2015). **

81. Tiihonen, J. *et al.* Genetic background of extreme violent behavior. *Molecular Psychiatry* **20**, 786–792 (2015). **

82. Aebi, M. *et al.* Gene-set and multivariate genome-wide association analysis of oppositional defiant behavior subtypes in attention-deficit/hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* **171**, 573–588 (2016). **

83. Brevik, E. J. *et al.* Genome-wide analyses of aggressiveness in attention-deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* **171**, 733–747 (2016). **

84. Merjonen, P. *et al.* Hostility in adolescents and adults: a genome-wide association study of the Young Finns. *Transl Psychiatry* **1**, e11 (2011). **

85. Mick, E. *et al.* Genome-wide association study of the child behavior checklist dysregulation profile. *J Am Acad Child Adolesc Psychiatry* **50**, 807–817 (2011). **

86. Mick, E. *et al.* Genome-Wide Association Study of Proneness to Anger. *PLoS One* **9**, 1–7 (2014). **

87. McGue, M. *et al.* A Genome-Wide Association Study of Behavioral Disinhibition. *Behavior Genetics* **43**, 363–373 (2013). **

88. Pappa, I. *et al.* A genome-wide approach to children's aggressive behavior: The EAGLE consortium. *Am J Med Genet B Neuropsychiatr Genet* **171**, 562–572 (2016). **

89. Ip, H. F. *et al.* Genetic association study of childhood aggression across raters, instruments and age. *Transl Psychiatry* **11**, 413 (2021). **

90. Van der Laan, C. *et al.* Continuity of genetic risk for aggressive behavior across the life-course. *Behavioral Genetics* **51** (5):592-606 (2021)

91. Barnes, J. C. *et al.* The propensity for aggressive behavior and lifetime incarceration risk: A test for gene-environment interaction (G × E) using whole-genome data. *Aggression and Violent Behavior* **49**, 101307 (2019).

92. Bonder, M. J. *et al.* Disease variants alter transcription factor levels and methylation of their binding sites. *Nature genetics* **49**, 131–138 (2017).

93. Gescher, D. M. *et al.* Epigenetics in Personality Disorders: Today's Insights. *Front Psychiatry* **9**, 579 (2018).*

94. Cecil, C. A. M. *et al.* DRD4 methylation as a potential biomarker for physical aggression: An epigenome-wide, cross-tissue investigation. *Am J Med Genet B Neuropsychiatr Genet* **177**, 746–764 (2018). ***

95. Tylee, D. S., Kawaguchi, D. M. & Glatt, S. J. On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics* **162B**, 595–603 (2013).

96. Smith, A. K. *et al.* DNA extracted from saliva for methylation studies of psychiatric traits: Evidence tissue specificity and relatedness to brain. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **168**, 36–44 (2015).

97. Barker, E. D. *et al.* A Methylome-Wide Association Study of Trajectories of Oppositional Defiant Behaviors and Biological Overlap With Attention Deficit Hyperactivity Disorder. *Child development* **89**, 1839–1855 (2018).

98. Philibert, R. A. *et al.* Gene environment interactions with a novel variable Monoamine Oxidase A transcriptional enhancer are associated with antisocial personality disorder. *Biological psychology* **87**, 366–371 (2011).

99. Wang, D. *et al.* Peripheral SLC6A4 DNA methylation is associated with in vivo measures of human brain serotonin synthesis and childhood physical aggression. *PloS one* **7**, e39501 (2012).

100. Provencal, N. *et al.* Differential DNA methylation regions in cytokine and transcription factor genomic loci associate with childhood physical aggression. *PLoS One* **8**, e71691 (2013).

101. Provencal, N. *et al.* Association of childhood chronic physical aggression with a DNA methylation signature in adult human T cells. *PLoS One* **9**, e89839 (2014).

102. Guillemin, C. *et al.* DNA methylation signature of childhood chronic physical aggression in T cells of both men and women. *PloS one* **9**, e86822 (2014).

103. Montalvo-Ortiz, J. L., Zhang, H., Chen, C., Liu, C. & Coccaro, E. F. Genome-Wide DNA Methylation Changes Associated with Intermittent Explosive Disorder: A Gene-Based Functional Enrichment Analysis. *Int J Neuropsychopharmacol* **21**, 12–20 (2018). ***

104. Montalvo-Ortiz, J. *et al.* M79 - Genome-wide DNA methylation signatures of aggression risk in children. *European Neuropsychopharmacology* **29**, S997–S998 (2019). ***

105. Van Dongen, J. *et al.* Epigenome-Wide Association Study of Aggressive Behavior. *Twin Research and Human Genetics* **18**, 686–698 (2015). ***

106. Van Dongen, J. *et al.* DNA methylation signatures of aggression and closely related constructs: A meta-analysis of epigenome-wide studies across the lifespan. *Molecular psychiatry* **26**, 2148–2162 (2021). ***

107. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).

108. Hagenbeek, F. A. *et al.* Urinary Amine and Organic Acid Metabolites Evaluated as Markers for Childhood Aggression: The ACTION Biomarker Study. *Frontiers in Psychiatry* **11**, (2020).

109. Whipp, A. M. *et al.* Ketone body 3-hydroxybutyrate as a biomarker of aggression. *Scientific Reports* **11**, 5813 (2021).

110. Tiihonen, J. *et al.* Neurobiological roots of psychopathy. *Molecular Psychiatry* **25**, 3432–3441 (2020).

111. Chen, X. *et al.* Profiling the differences of gut microbial structure between schizophrenia patients with and without violent behaviors based on 16S rRNA gene sequencing. *International Journal of Legal Medicine* **135**, 131–141 (2021).

112. Cohen, P. The origins of protein phosphorylation. *Nature Cell Biology* **4**, (2002).

113. Miles, D. R. & Carey, G. Genetic and environmental architecture on human aggression. *Journal of Personality and Social Psychology* **72**, 207–217 (1997).

114. Rhee, S. H. & Waldman, I. D. Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin* **128**, 490–529 (2002).

115. Burt, S. A., Klump, K. L., Gorman-Smith, D. & Neiderhiser, J. M. Neighborhood Disadvantage Alters the Origins of Children's Nonaggressive Conduct Problems. *Clinical Psychological Science* **4**, 511–526 (2016).

116. CONVERGE consortium *et al.* Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588–591 (2015).

117. Vassos, E., Collier, D. A. & Fazel, S. Systematic meta-analyses and field synopsis of genetic association studies of violence and aggression. *Mol Psychiatry* **19**, 471–477 (2014).

118. Ormel, J., Hartman, C. A. & Snieder, H. The genetics of depression: successful genome-wide association studies introduce new challenges. *Translational Psychiatry* **9**, 114 (2019).

119. Holland, D. *et al.* Beyond SNP Heritability: Polygenicity and Discoverability of Phenotypes Estimated with a Univariate Gaussian Mixture Model. *bioRxiv* 133132 (2019) doi:10.1101/133132.

120. Kukekova, A. v, Temnykh, S. v, Johnson, J. L., Trut, L. N. & Acland, G. M. Genetics of behavior in the silver fox. *Mammalian genome : official journal of the International Mammalian Genome Society* **23**, 164–177 (2012).

121. Hecht, E. E. *et al.* Neuromorphological changes following selection for tameness and aggression in the Russian fox-farm experiment. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **41**, 6144–6156 (2021).

122. Takahashi, A. & Miczek, K. A. Neurogenetics of aggressive behavior: studies in rodents. *Curr Top Behav Neurosci* **17**, 3–44 (2014).

123. Gouveia, F. V. *et al.* Amygdala and Hypothalamus: Historical Overview With Focus on Aggression. *Neurosurgery* **85**, 11–30 (2019).

124. Rosenfeld, C. S. *et al.* Hypothalamic transcriptome of tame and aggressive silver foxes (Vulpes vulpes) identifies gene expression differences shared across brain regions. *Genes, brain, and behavior* **19**, e12614 (2020).

125. Wörheide, M. A., Krumsiek, J., Kastenmüller, G. & Arnold, M. Multi-omics integration in biomedical research – A metabolomics-centric review. *Analytica Chimica Acta* **1141**, 144–162 (2021).

126. Wörheide, M. A., Krumsiek, J., Kastenmüller, G. & Arnold, M. Multi-omics integration in biomedical research – A metabolomics-centric review. *Analytica Chimica Acta* **1141**, 144–162 (2021).

127. eGTAxProject. Enhancing GTEx by bridging the gaps between genotype, gene expression, and disease. *Nature Genetics* **49**, 1664–1670 (2017).

128. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genetics* **48**, 481–487 (2016).

129. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nature Genetics* **48**, 245–252 (2016).

130. Goh, K.-I. et al. The human disease network. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 8685–90 (2007).

131. Lowe, W. L., Reddy, T. E. & Reddy, T. E. Genomic approaches for understanding the genetics of complex disease. *Genome research* **25**, 1432–41 (2015).

132. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical pharmacology and therapeutics* **69**, 89–95 (2001).

133. Khabsa, M., Elmagarmid, A., Ilyas, I., Hammady, H. & Ouzzani, M. Learning to identify relevant studies for systematic reviews using random forest and external information. *Machine Learning* **102**, 465–482 (2016).

134. Cohen, A. M., Hersh, W. R., Peterson, K. & Yen, P.-Y. Reducing Workload in Systematic Review Preparation Using Automated Citation Classification. *Journal of the American Medical Informatics Association* **13**, 206–219 (2006).

135. Borah, R., Brown, A. W., Capers, P. L. & Kaiser, K. A. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ open* **7**, e012545 (2017).

136. Chen, X. et al. Dysregulation of amino acids and lipids metabolism in schizophrenia with violence. *BMC Psychiatry* **20**, 1–11 (2020).

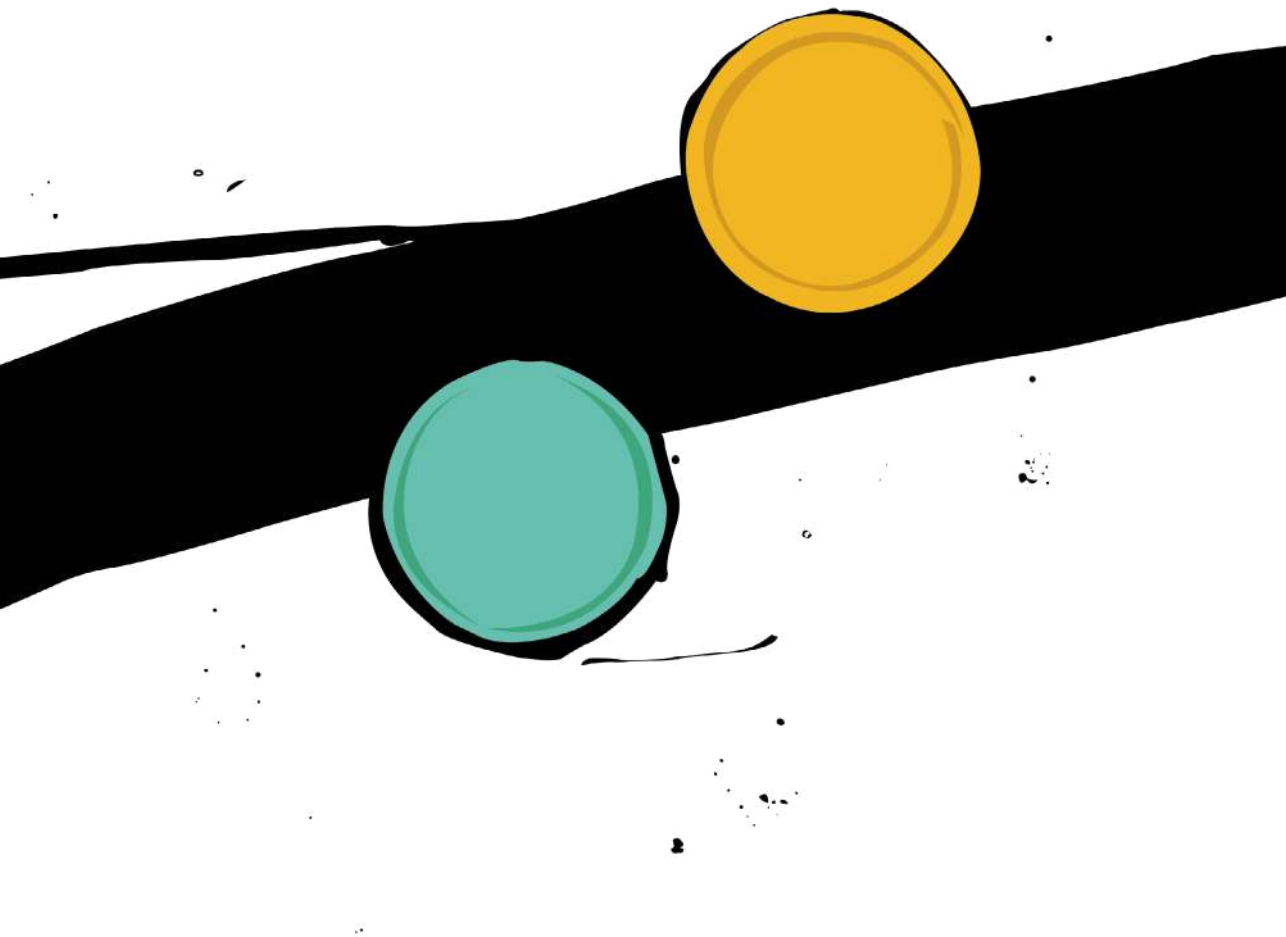
Supplements

The supplementary materials are available online

https://journals.lww.com/psychgenetics/Fulltext/2019/10000/Genomics_of_human_aggression__current_state_of.6.aspx

Chapter 8

Predicting complex traits and exposures from polygenic scores and blood and buccal DNA methylation profiles



Abstract

We examined the performance of methylation scores (MS) and polygenic scores (PGS) for birth weight, BMI, prenatal maternal smoking exposure, and smoking status to assess the extent to which MS could predict these traits and exposures over and above the PGS in a multi-omics prediction model. Methylation scores may be seen as the epigenetic equivalent of PGS, but because of their dynamic nature and sensitivity of non-genetic exposures may add to complex trait prediction independently of PGS. MS and PGS were calculated based on genotype data and DNA-methylation data in blood samples from adults (Illumina 450K; *N*=2,431; mean age 35.6) and in buccal samples from children (Illumina EPIC; *N*=1,128; mean age 9.6) from the Netherlands Twin Register. Weights to construct the scores were obtained from results of large epigenome-wide association studies (EWASs) based on whole blood or cord blood methylation data and genome-wide association studies (GWASs).

In adults, MSs in blood predicted independently from PGSs, and outperformed PGSs for BMI, prenatal maternal smoking, and smoking status, but not for birth weight. The largest amount of variance explained by the multi-omics prediction model was for current vs never smoking (54.6%) of which 54.4% was captured by the MS. The two predictors captured 16% of former vs never smoking initiation variance (MS:15.5%, PGS: 0.5%), 17.7% of prenatal maternal smoking exposure variance (MS:16.9%, PGS: 0.8%), 11.9% of BMI variance (MS: 6.4%, PGS 5.5%) and 1.9% of birth weight variance (MS: 0.4%, PGS: 1.5%).

In children, MSs in buccal samples did not show independent predictive value. The largest amount of variance explained by the two predictors was for prenatal maternal smoking exposure (2.6%), where the MSs contributed 1.5%.

These results demonstrate that blood DNA MS in adults explain substantial variance in current smoking, large variance in former smoking, prenatal smoking, and BMI, but not in birth weight. Buccal cell DNA methylation scores have lower predictive value, which could be due to different tissues in the EWAS discovery studies and target sample, as well as to different ages. This study illustrates the value of combining polygenic scores with information from methylation data for complex traits and exposure prediction.

Published as: Odintsova, V. V., Rebattu, V., Hagenbeek, F. A., Pool, R., Beck, J. J., Ehli, E. A., van Beijsterveldt, C., Ligthart, L., Willemsen, G., de Geus, E., Hottenga, J. J., Boomsma, D. I., van Dongen, J. (2021) Predicting Complex Traits and Exposures From Polygenic Scores and Blood and Buccal DNA Methylation Profiles. *Frontiers in Psychiatry,* 12:688464.

Introduction

Nearly all complex traits in humans are a function of their genotype and of environmental exposures, as shown by family and twin studies[1–3]. DNA-based predictors of complex traits can increasingly serve to improve prediction of health outcomes and disease and to optimize risk stratification[4] and are also considered for application in social sciences and education[5,6]. Whereas, DNA-based predictors are static and solely capture genomic information, other predictors such as those based on epigenome data are dynamic and may capture both genetic and environmental information.

Polygenic scores (PGS; sometimes referred to as Polygenic Risk Scores) are defined as the weighted sum of an individual's risk alleles, or increasing alleles for a continuous trait, of a pre-selected number of single nucleotide polymorphisms (SNPs). In some areas of medicine, polygenic risk scores are already beginning to be employed to predict individual risk of disease[7–9]. The PGS of an individual for a trait is calculated by multiplying, for each SNP, the number of risk alleles by a weight and then summing over all SNPs. Weights are typically estimated in a regression analysis, from a genome-wide association study (GWAS) for the trait from an independent discovery sample (typically, a large GWAS meta-analysis), and are included in the GWAS summary statistics (i.e., the estimated effect sizes, the standard errors of the estimates and the corresponding p-values).

This polygenic type of approach can be generalized to other omics data, including epigenomics where it results in DNA methylation scores (MS)[10], which can be described as weighted sums of the individual's methylation levels of a selected number of CpG sites. The individual's methylation levels at each CpG in an independent study population are multiplied by their corresponding weights and summed over multiple sites. Here the weights are based on summary statistics from a single or a meta-analysis epigenome-wide association study (EWAS) of the trait. By combining the effects of multiple CpG sites into a MS, a larger proportion of variance in traits is likely be explained compared to the variance that is captured by individual CpG sites. In addition to their value for prediction of complex traits and disease risk, MSs could potentially be informative as biomarkers for environmental exposures[11] or to monitor disease progression, and might be considered in association analyses in which individual CpG sites do not achieve significance or as a dimension reduction approach in interaction and mediation analyses[12,13].

The number of genetic variants and CpG sites associated with complex traits is growing based on findings from GWAS and EWAS meta-analyses. Birth weight was associated with 60 independent signals in a multi-ancestry GWA meta-analysis, capturing up to 4.9% of the variance in birth weight in different cohorts[14], and with 914 epigenome-wide Bonferroni-significant CpGs in an EWAS meta-analysis of multiple birth cohorts with cord blood DNA methylation data[15]. Body mass index (BMI) was associated with 751 SNPs in adults in the currently largest European ancestry GWAS meta-analysis, capturing approximately 6% of the BMI variance[16]. The currently largest EWAS meta-analysis of BMI based on whole blood from adults identified association with 278 Bonferroni-significant CpGs[12]. Smoking initiation was associated with 566 genetic variants in a GWAS of more than one million individuals, capturing 3.6% and 4.2% of the variance in the trait in prediction cohorts[17]. A large EWAS meta-analysis of smoking identified 18,760 CpGs

significantly differentially methylated in relation to current smoking in adults at a false discovery rate (FDR) of 5% from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium, and 2,623 FDR significant CpGs in association with former smoking[18]. EWAS meta-analyses conducted in newborns using cord blood DNA methylation data identified 6,073 CpGs with FDR significance in association with prenatal maternal smoking[19].

Attempting to capture the DNA methylation differences, previous studies have developed polygenic methylation predictors. We extensively reviewed the literature on studies that report methylation predictors as single MS and studies that examined the combined predictive value of MS and PGS (see **Supplementary Table S1**). Taking the results from EWASs into independent target samples in which MSs are defined, has yielded promising results for birth weight[20], BMI[20–22], prenatal maternal smoking[23,24], and smoking status[11,23,25–28]. Reed and colleagues[20] computed MSs for birth weight based on the 135 CpGs from an adult BMI EWAS in the Framingham Heart Study and Lothian Birth cohorts (*N*=3,743)[29]. These scores captured 2% of birth weight variation in 823 ALSPAC newborns with DNA methylation data in cord blood, which was higher than the variance captured by a PGS (0.4%). Several studies created whole blood DNA MSs of BMI and made predictions in children and adults. MSs based on 78 probes from 2,377 adults of the Framingham Heart Study and weights (effect sizes) from 750 adults of the LifeLines DEEP study explained 11% of the variance in BMI in 1,366 adults from Lothian Birth cohorts and 5% of BMI variance in 403 adolescents from Brisbane Systems Genetic Study (BSGS). MSs based on 400 CpGs from 2,562 Generation Scotland participants explained 10% of BMI variance in 892 adults from Lothian Birth cohort[21]. MSs based on 135 probes from 3,742 adults from both Framingham Heart Study and Lothian Birth cohorts explained 10% of BMI variance in 726 ALSPAC women and up to 3% of BMI variance in children at different ages[20]. It has been shown that MS for BMI perform better in adults compared to children and adolescents[20,22]. Attempts of cross-tissue performance testing were scarce[25,30], however, it have been shown that some alterations persist across tissue types[31].

For prenatal maternal smoking, MS based on weights from cord blood DNA methylation EWASs of 1,057 newborns from Norwegian Mother and Child Cohort Study (MoBa) was tested on another MoBa subset of 221 newborns[24], and MS based on weights from cord blood DNA methylation EWAS meta-analysis of 6,685 newborns done by Joubert and colleagues[19] was tested on 754 ALSPAC women around 30 years old[23]; the predictive accuracy (the amount of variation in the outcome explained by the score) was lower in women than in newborns. Smoking predictors have been described based on different numbers of probes from whole blood DNA methylation studies. Only 2 CpGs were included the smoking MS of Zhang et al. that predicted smoking status in 9,949 older adults[28]. The largest smoking MSs included 2,623 Bonferroni significant CpGs from EWAS meta-analysis of 15,907 individuals[18] and predicted smoking status during pregnancy in 754 women by Richmond et al.[23]. The same CpGs were used by Sugden et al.[11] to predict smoking status in 1,037 adults from the Dunedin Longitudinal Study and 2,232 twins from the Environmental Risk Longitudinal Study.

Despite the growing number of cohorts that have both genomic and methylation data, few attempts have been made to combine PGS and MS in a multi-omics model. To the best

of our knowledge, BMI and height are currently the only traits for which the prediction by PGS and MS combined has been investigated [21,22]. In a combined model, the PGS and MS together explained 17% of the variance in BMI in 1,366 adults[22] and 18% in 889 adults[21], both from the Lothian Birth cohorts, 13–16% in 750 adults from Lifelines and 8% in adolescents from the Brisbane Systems Genetic Study 32, corresponding to an added ~4–9% extra variance explained compared to the PGS alone.

We expand on the previous work by addressing several points. First, it is largely unknown to what extent MS based on EWAS weights derived in adults predict trait variation in children and vice versa. Second, previous studies of MS were based on cord blood or whole blood, and it is unknown if these scores translate to other tissues. Third, for all traits, except BMI and height[20,22], it is unknown whether MS add to prediction independently of PGS.

In the current biomarker study, we analyze the predictive accuracy of PGS and MS (both individually and combined). The goal of our study is to examine if the MSs add predictive value above the PGSs. The weights required for DNA methylation data were obtained large EWAS and applied to methylation levels from two different tissues (blood and buccal). We analyze data from large groups of adults with DNA methylation in blood (*N* = 2,431, mean age = 35.6) and children with DNA methylation in buccal cells (*N* = 1,128, mean age = 9.6) who participate in research projects of the Netherlands Twin Register and consider multiple traits. For an early-life trait we analyze birth weight, and for a trait that is dynamic in childhood and adulthood, we analyze BMI. As early and later life exposures we examine prenatal maternal smoking during pregnancy and own smoking. These four phenotypes represent complex traits and exposures with different relative contributions of genetics and environment to inter-individual variance.

Materials and Methods

Overview

This study included adults and children who participated in studies from the Netherlands Twin Register (NTR). DNA samples in adult twins and family members were isolated from whole blood DNA data and in twin children from buccal cells. Adults took part in the NTR-Biobank[32] and children in the FP7-Action project[33–36]. The study was approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Centre, Amsterdam, an Institutional Review Board certified by the U.S. Office of Human Research Protections (IRB number IRB00002991 under Federal-wide Assurance FWA00017598; IRB/institute codes, NTR 03-180). Adults provided written informed consent, for children consent was given by their parents.

Adults

Study population and samples

After quality control, genome-wide DNA methylation profiles in whole blood and genotype data were available for 2,431 NTR adults[37]. This dataset included 2,426 individuals from twin pairs, and 5 family members (mothers and spouses). The mean age at DNA collection was 35.6 years (range = 17.6–79.2 years) and 32.7% of subjects were males. For 20 participants, longitudinal methylation data (methylation data at two time points) were available. Individuals with missing data on phenotypes or covariates, and phenotype outliers were excluded from analysis, resulting in a sample size of 2,040 for birth weight, 2,410 for BMI, 1,914 for current vs never smoking, and 1,938 for former vs never smoking. Because prenatal maternal smoking exposure is equal for co-twins, one twin from each pair was randomly included in the analysis, resulting in a sample size of 720. The blood sampling procedure has been described by Willemsen et al.[32].

DNA methylation

DNA methylation in blood was assessed with the Infinium HumanMethylation450 BeadChip Kit (Illumina, San Diego, CA, USA) by the Human Genotyping facility (HugeF) of ErasmusMC, the Netherlands (http://www.glimdna.org/) as part of the Biobank-based Integrative Omics Study (BIOS) consortium[38]. DNA methylation measurements have been described previously[37,38]. Genomic DNA (500ng) from whole blood was bisulfite treated using the Zymo EZ DNA Methylation kit (Zymo Research Corp, Irvine, CA, USA), 12μl of buffer was utilized to elute the converted DNA off the column after conversion, and 4 μl (~33ng/μl) of bisulfite-converted DNA was measured on the Illumina 450K array following the manufacturer's protocol. A number of sample- and probe-level quality checks and sample identity checks were performed, as described in detail previously[37]. In short, sample-level QC was performed using MethylAid[39]. Probes were set to missing in a sample if they had an intensity value of exactly zero, or a detection $p > 0.01$, or a bead count of < 3. After these steps, probes that failed based on the above criteria in > 5% of the samples were excluded from all samples (only probes with a success rate ≥ 0.95 were retained). For all samples, ambiguously mapped probes were excluded, based on the definition of an overlap of at least 47 bases per probe from Chen et al.[40], and all probes containing a SNP, identified in the Dutch population[41], within the CpG site (at the C or G position) were excluded, irrespective of minor allele frequency. Only autosomal sites were kept in the current analyses (*N*=411,169). The methylation data were normalized with functional normalization[42]. Probes with missing values (probes with missing values in more than 5% of the sample were removed) were imputed with the function imputePCA from the package missMDA as implemented in the pipeline for DNA methylation array analysis developed by the Biobank-based Integrative Omics Study (BIOS) consortium[43].

Phenotyping

Data on birth weight were obtained from self-report or by parental report. If data were available from multiple surveys by Adult Netherlands Twin Register (ANTR) and/or informants, they were checked for consistency[44]. When multiple data points differed by less than 200 grams, the average was taken, and in the cases of larger differences,

data were excluded. Information on maternal smoking during pregnancy was obtained in ANTR Survey 10 (data collection in 2013) with the following question: “Did your mother ever smoke during pregnancy?” Answer categories were “no”, “yes” and “I don’t know”. For twin pairs, the answers were checked for consistency and missing data for one twin were supplemented with data from the co-twin where possible. In the case of inconsistent answers, the data from both co-twins were set to missing. If both twins answered “I don’t know”, the variable was coded as missing. Data on body mass index (BMI) and smoking status were collected at blood draw[32]. We analyzed two smoking phenotypes: current smokers (1) versus never smokers (0), and former smokers (1) versus never smokers (0). The percentage of white blood cell was obtained in fresh blood samples collected in EDTA (Ethylene Diamine Tetra Acetic acid) tubes[45]. For birth weight and BMI, we removed outliers using a cut-off of 3 standard deviations from the mean. For birth weight, 6 outliers were removed; for BMI, 27 outliers were removed.

Children

Study population and samples

Genotype data and genome-wide DNA methylation profiles in buccal cells were collected in a children that participated in a larger project on childhood aggression “Aggression in Children: Unraveling gene-environment interplay to inform Treatment and InterventiON strategies” (ACTION; http://www.action-euproject.eu/) and consists of twins who score high or low on aggression[33–36]. After quality control, genome-wide DNA methylation data and genotype data were available for 1,128 children from twin pairs (mainly monozygotic twins). The mean age at DNA collection was 9.6 years (range = 5.6–12.9 years) and 52.8% were males. For 2 participants, a technical replicate measure on with the Infinium MethylationEPIC BeadChip Kit was included[36]. Individuals without missing data on phenotypes or covariates were included in the analyses, and phenotype outliers were excluded, resulting in a sample size of 1,070 children for birth weight and 1,072 for BMI. Because prenatal maternal smoking exposure is equal for co-twins, one twin from each pair was randomly included in the analysis, resulting in a sample size of 547. The sample collection protocol is available at: http://www.action-euproject.eu/content/data-protocols. DNA was collected from buccal swabs at home: 16 cotton sticks were individually rubbed against the inside of the cheek in the morning and evening on 2 days by the participants and placed in buffer. Individuals were asked to refrain from eating or drinking 1 hour prior to sampling. High molecular weight genomic DNA was extracted from the swabs by standard DNA extraction techniques and visualized using agarose gel electrophoresis. The DNA samples were quantified using the Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific, Waltham, MA, USA).

DNA methylation

DNA methylation was assessed with the Infinium MethylationEPIC BeadChip Kit (Illumina, San Diego, CA, USA) by the Human Genotyping facility (HugeF) of ErasmusMC, the Netherlands (http://www.glimdna.org/)[36]. Quality control (QC) and normalization of the methylation data were performed using a pipeline developed by the Biobank-based

Integrative Omics Study (BIOS) consortium[43], which includes sample quality control using the R package MethylAid[39] and probe filtering and functional normalization[42] as implemented in the R package DNAmArray. The following probe filters were applied: probes were set to missing (NA) in a sample if they had an intensity value of exactly zero, detection *P*-value > 0.01, or bead count < 3; probes were excluded from all samples if they mapped to multiple locations in the genome, if they overlapped with a SNP or Insertion/Deletion (INDEL), or if they had a success rate < 0.95 across samples. Annotations of ambiguous mapping probes (based on an overlap of at least 47 bases per probe) and probes where genetic variants (SNPs or INDELS) with a minor allele frequency > 0.01 in Europeans overlap with the targeted CpG or single base extension site (SBE) were obtained from (Pidsley et al. 2016). For two twins, a technical replicate measure on EPIC was obtained (on different BeadChip Arrays). Probes with missing values (probes with missing values in more than 5% of the sample were removed) were imputed with the function imputePCA from the package missMDA as implemented in the pipeline for DNA methylation array analysis developed by the BIOS consortium[43].

Phenotyping

Data on birth weight of the young twins came from surveys sent to mothers shortly after the registration of the newborn twins[47]. Data on BMI were collected from surveys filled out by mothers and fathers in the Young Netherlands Twin Register (YNTR) when children were around 5, 7, 10 and 12 years of age. If both parents completed the survey, preference was given to data provided by the mother. BMI closest to the date of DNA collection was selected. The average time between DNA collection and BMI assessment was 1.9 years before the survey (median=-0.9, range: from buccal sample collection 10.3 years before survey to buccal sample collection 2.1 years after survey). Information on maternal smoking during pregnancy was reported by mothers after registration for three trimesters of pregnancy and was coded as “non-smoking” if the mother did not smoke during the entire pregnancy and “smoking” if the mother smoked at least during one trimester[48]. For birth weight and BMI, we removed outliers using a cut-off of 3 standard deviations from the mean. For birth weight, 1 outlier was removed; for BMI, 12 outliers were removed.

Cellular proportions were predicted with hierarchical epigenetic dissection of intra-sample-heterogeneity (HepiDISH) with the RPC method (reduced partial correlation), as described by[49] and implemented in the R package HepiDISH. HepiDISH is a cell-type deconvolution algorithm developed for estimating cellular proportions in epithelial tissues based on genome-wide methylation profiles and makes use of reference DNA methylation data from epithelial cells, fibroblasts and seven leukocyte subtypes. This method was applied to the data after data QC and normalization.

Genotyping

Genotyping in children (YNTR) and adults (ANTR) was done on multiple platforms over time including Perlegen-Affymetrix, Affymetrix 6.0, Affymetrix Axiom, Illumina Human Quad Bead 660, Illumina Omni 1M and Illumina GSA. Quality control and processing of the genotype data was performed on the complete dataset of all genotyped participants from the NTR. Quality control was carried out and haplotypes were estimated in PLINK. CEU population outliers, based on per platform 1000 Genomes PC projection with the

TABLE 1.
Discovery epigenome-wide and genome-wide association studies that provided the summary statistics to calculate DNA methylation scores and polygenic scores

Trait/ Exposure	Reference	Phenotype in discovery study –DNA methylation Tissue	*N* discovery cohort used to create scores	*N* CpGs/SNPs reported at significant level in reference
EWAS				
Birth weight	Küpers et al.[15]	Birth weight – Cord blood	8,825 newborns	914 (Bonferroni)
BMI	Wahl et al.[12]	BMI – Whole blood	5,387 adults	5,547 (FDR <0.05)
Prenatal Maternal Smoking	Sikdar et al.[51]*	Prenatal maternal smoking – Cord blood	4,994 newborns (897 exposed to sustained prenatal smoking)	34,541 (FDR <0.05)
Smoking	Sikdar et al.[51]*	Current vs. never smoked – Whole blood	9,389 adults (2,433 current smokers)	2,623 (FDR <0.05)
	Joehanes et al, 2016[18]	Former vs. never smoked – Whole blood	13,474 adults (6,518 former smokers)	
GWAS				
Birth weight	http://www.nealelab.is/uk-biobank/	Birth weight	280,250 (UK biobank)	not published
BMI	Yengo et al.[16]	BMI	~700,000 individuals of different ancestry 652,099** were used in our study for calculation of PGS	941
Prenatal Maternal Smoking	UK Biobank http://www.nealelab.is/uk-biobank/	Maternal smoking around birth	331,862 (UK Biobank)	not published
Smoking	Liu et al.[17]	Smoking initiation (ever/never smoked)	Up to 1.2 million individuals in discovery study 625,536** used in our study for PGS	566

EWAS, Epigenome-wide association study; GWAS, Genome-wide association study; FDR, False Discovery Rate.
* Sikdar et al.[51] repeated the meta-analysis by Joubert et al.[19] (EWAS in newborns) and Joehanes et al.[18] (EWAS in never vs current smokers), and provided full genome-wide summary statistics for a fixed-effects meta-analysis of maternal smoking in newborns (cord blood) and for current versus never smokers (whole blood).
** NTR and 23andMe are excluded. For additional details, see **Supplementary Table 1**.

Smartpca software[50], were excluded. Data were phased per platform using Eagle, and then imputed to 1000 Genomes using Minimac, following the Michigan imputation server protocols. For the polygenic scoring imputed

data were converted to best guess genotypes, and filtered to include only ACGT SNPs, SNPs with MAF > 0.01, HWE $p > 10^{-5}$ and genotype call rate > 0.98, and exclude SNPs with more than 2 alleles. All Mendelian errors were set to missing. Principal components (PCs) were calculated with Smartpca using linkage-disequilibrium-pruned (LD-pruned) 1000 Genomes–imputed SNPs that were also genotyped on at least one platform, had MAF > 0.05 and were not present in the long-range LD regions.

Statistical methods

EWAS and GWAS summary statistics

MSs and PGSs were created using weights based on large epigenome-wide association study (EWAS) and genome-wide association study (GWAS) meta-analyses. These studies are summarized in **Table 1**. Additional information on the studies and derived scores is provided in **Supplementary Tables S2- S3**.

DNA methylation scores

The effect sizes obtained from the summary statistics from previously published EWAS meta-analyses (**Table 1** and **Supplementary Table S1**) were used to calculate weighted MSs in NTR participants as previously done by Elliot et al.[26], Shah et al.[22], Wahl et al.[12], Richmond et al.[23], and Sugden et al.[11]. For each trait and for each individual, a score was calculated by multiplying the methylation level at a given CpG by the previously reported effect size of the CpG (β), and then summing these values over all CpGs:

$$\textit{DNA methylation score} = \beta_1 * CpG_1 + \beta_2 * CpG_2 \ldots. + \beta_i * CpG_i$$

where CpG_i is the methylation level at CpG site *i*, which ranges between 0 and 1, and β_i is the effect size (regression coefficient) at the CpG_i obtained from summary statistics of EWAS meta-analyses that did not include participants from the NTR.

For each phenotype, except for former smoking, we calculated multiple MSs based on different subsets of CpGs according to their significance level. Subsets of CpGs were selected based on p-value thresholds of $<1x10^{-1}$, $<1x10^{-5}$, and $<1x10^{-7}$. For former versus never smoking, genome-wide summary statistics were not available, and we calculated MSs for former smoking based on CpGs that were significant in the EWAS of former versus never smokers at a False Discovery Rate of 5%. Additionally, we tested prediction of former smoking based on the MSs derived from the genome-wide EWAS summary statistics of current versus never smoking. To examine if removal of CpGs with correlated DNA methylation levels affects trait prediction, we also calculated pruned scores by step-wise selection of the most significant CpG site and excluding CpG sites with a correlation of 0.1 or higher (threshold chosen based on Shah et al.[22]) in order to keep an independent set of CpGs.

Polygenic scores

Polygenic scores (PGSs) were calculated based on weighting of genotypes by effect sizes as made available from GWAS summary statistics (see **Table 1** and **Supplementary Table S2**) in discovery samples without NTR. Before calculating the PGSs, linkage disequilibrium (LD) weighted β's were calculated from these summary statistics by the LDpred package to correct for the effects of LD and to maximize predictive accuracy of the PGSs[52]. QC has been applied: MAF>0.01, duplicated SNPs, mismatching alleles, ambiguous SNPs were excluded. We randomly selected 2500 unrelated individuals from NTR as a reference population to calculate the LD patterns. The adjusted β's were calculated from an LD pruning window of 250 KB, with the fraction of causal SNPs set at 0.01 for birth weight, because this fraction was previously shown to perform optimally for birth weight in the NTR population[53] and at 0.50 for other phenotypes. The PGSs were obtained for all NTR participants with genotyping data with the PLINK 1.9 software.

Statistical analysis

Continuous traits, MSs and PGSs were z-score transformed [trait value – trait mean/ trait standard deviation] before analysis. Pairwise Pearson correlations between each trait, MSs, PGSs and covariates were computed in NTR adults and children EWAS datasets for each phenotype and visualized in correlation plots. For each trait, we fitted a series of regression models to examine: (1) the predictive value of MSs; (2) the predictive value of a PGS; and (3) whether MS and PGS contributed independently to trait prediction in a combined predictor. First, for each trait, we evaluated the performance of multiple different MSs based on different p-value thresholds, pruned and unpruned. We took the score that explained the largest amount of variance forward to the combined model. Second, we evaluated the performance of PGS in prediction of each trait. Finally, we examined if MSs predict these traits over and above the PGSs and estimated how much variance in each trait was explained by multi-omics predictor, e.g. by MSs and PGSs together. Sex and age at DNA collection were included as covariates in all three models. In the prediction models with whole blood DNA MSs, we corrected for percentages of neutrophils, monocytes and eosinophils. In the prediction models with buccal DNA MSs, we corrected for epithelial cell and natural killer cell proportions. To adjust for technical variation, array row and bisulfite plate (dummy-coding) were included as covariates in all models with EWAS covariates. In models including PGSs, we corrected for genotype data-specific covariates: the first ten genetic principal components and genotype platform dummy variables (GWAS covariates).

Continuous traits. For birth weight and BMI, the following models were fitted in each of the two datasets (whole blood methylation data from adults and buccal methylation data from children):

Model 1: Trait ~ MS + sex + age + EWAS covariates

Model 2: Trait ~ PGS + sex + age + GWAS covariates

Model 3: Trait ~ MS + PGS + sex + age + EWAS covariates + GWAS covariates

Analyses were carried out with generalized estimation equation (GEE) models accounting for familial relatedness, fitted with the R package 'gee' with the following settings: Gaussian link function for continuous data, 100 iterations, and the 'exchangeable' option to account for the correlation structure within families. To calculate the variance explained by the MS and the PGS, we squared the regression coefficient of each score obtained in GEE. This value was multiplied by 100 to obtain the percentage of variance explained.

Dichotomous traits. For dichotomous traits, i.e. prenatal maternal smoking, current vs never smoking, and former vs never smoking, the following models were fitted in two datasets (whole blood methylation data from adults and buccal methylation data from children):

Model 1a: Trait ~ MS + sex + age + EWAS covariates

Model 1b: Trait ~ sex + age + EWAS covariates

Model 2a: Trait ~ PGS + sex + age + GWAS covariates

Model 2b: Trait ~ sex + age + GWAS covariates

Model 3a: Trait ~ MS + PGS + sex + age + EWAS covariates + GWAS covariates

Model 3b: Trait ~ PGS + sex + age + EWAS covariates + GWAS covariates

Model 3c: Trait ~ MS + sex + age + EWAS covariates + GWAS covariates

To obtain the variance explained, models were fitted with logistic regression with binomial family setting (link = "logit"). Estimation of the variance explained by the MS and PGS, was based on the approach proposed by Lee et al. where coefficients of determination (R^2) for binary responses are calculated on the liability scale[54]. R^2 is equal to the explained variance divided by the total variance; that is the sum of explained variance and residual (homoscedastic) variance. We first regressed the trait on the MS, sex, age and EWAS covariates (model 1a), and then on sex, age, and EWAS covariates only (model 1b). We calculated variance explained by all predictors in each model. We calculated the predictive value of the MS by subtracting the difference between the variance explained by the model 1a and 1b. The same was done for models with PGS with sex, age and GWAS covariates (model 2a and 2b), and then for combined model with both MS and PGS scores (models 3a-c). In the last case, the difference between explained variance in model 3a and model 3b gave us an estimate explained by MS, and the difference between explained variance in model 3a and model 3c resulted in estimate explained by PGS.

To correct for relatedness in smoking prediction, p-values were obtained from GEE models, fitted with the R package 'gee', with the binomial link function for dichotomous data, 100 iterations, and the 'exchangeable' option to account for the correlation structure within families. For prenatal smoking exposure (yes/no) we randomly chose one of the twins from the pair, and p-values were obtained from logistic regression models.

TABLE 2.
Characteristics: NTR adults and children

	Adults	Children
DNA methylation tissue	**Whole blood**	**Buccal cells**
DNA methylation array	Illumina 450k	Illumina EPIC
Sample size	2,431	1,128
Age, mean (sd)	35.6 (11.9)	9.6 (1.9)
Males, n (%)	794 (32.7%)	596 (52.8%)
Females, n (%)	1637 (67.3%)	532 (47.2%)
Birth weight, mean (sd)	2507.8 (573.7)	2395.9 (544.1)
missing	*404*	*59*
BMI, mean (sd)	24.1 (3.9)	16 (2.1)
missing	*14*	*46*
Prenatal maternal smoking (no), n (%)	1169 (90.1%)	921 (91.1%)
Prenatal maternal smoking (yes), n (%)	129 (9.9%)	90 (8.9%)
missing	*1133*	*117*
Never smokers, n (%)	1395 (57.5%)	NA
Former smokers, n (%)	527 (21.7%)	NA
Current smokers, n (%)	506 (20.8%)	NA
missing	*3*	NA
Cell counts, mean (sd)		
Neutrophil percentage	52.5 (9.1)	NA
Eosinophil percentage	3.1 (2.3)	NA
Monocyte percentage	8.4 (2.4)	NA
Epithelial cell proportion	NA	0.806 (0.116)
NK cell proportion	NA	0.03 (0.013)

NTR, Netherlands Twin Register; sd, standard deviation; NA, not available; NK, natural killer cells.
Descriptives are provided for covariates that were included as covariates in the prediction models.
The number of samples with complete data on DNA methylation, cell counts, and genotyping data are presented. For a small number of individuals, two samples are included in the analysis in adults (n = 20) and children (n = 2).

Sensitivity analysis

We carried out a sensitivity analysis in which we repeated the models for BMI prediction in children from MSs after removal of children for whom information on BMI was collected more than 3 years before or after DNA collection (N = 324 children removed; new N = 748).

Multiple testing correction

Statistical significance was assessed following Bonferroni correction for multiple testing (six tests in model 1 for birth weight, BMI, prenatal maternal smoking and current smoking,

seven tests in model 1 for former smoking, one test in models 2 and 3 for each trait in adults; the same number of tests in children except smoking status and plus eight tests in sensitivity analysis for BMI). This resulted in a significance level of 0.0012 (α = 0.05/42) for adults and 0.0016 (α = 0.05/32) for children.

Results

Characteristics of the NTR adult and children are presented in **Table 2**. Distribution of the traits/exposures as main outcomes and MSs and PGs as predictors are presented in **Supplementary Figures S1-S3**. Correlations between trait/expoosure, PGS, MSs, sex, age, and cellular compositions of the samples are shown in **Supplementary Figures S4-S5**. The correlations between PGS and MSs for the same trait were weak in adults (r = [0.01-0.15]) and children (r = [0.01-0.05]). Further, we report the correlation between the PGS and the MS that captured the largest amount of variation in the trait. We examined prediction of each phenotype by its MS and PGS separately. The explained variance and corresponding p-values for unpruned and pruned MSs with different thresholds for inclusion of CpGs are presented in **Table 3**, and for PGS in **Table 4**. To examine to what extent the PGS and the MS capture independent information, we fitted the model in which the outcome was regressed on both scores as multi-omics prediction presented in **Table 5. Figure 1** shows the variance explained by the MSs and PGSs separately and together as multi-omics predictor in previous and our studies.

Birth weight

The birthweight MSs were calculated based on the birthweight EWAS of cord blood samples from neonates[15]. The results of GEE showed that none of the MSs was strongly associated with birth weight in adults (p<0.0012) and children (p<0.0016). The pruned blood MS based on 934 CpGs with a p-value lower than $1x10^{-1}$ performed better in prediction of birth weight in adults compared with unpruned and other threshold pruned scores, accounting for 0.39% of the variance (p=0.004) (**Table 3**, **Figure 1A**). The PGS significantly predicted birth weight in adults (variance explained by PGS=1.52%, *p*=$1.56x10^{-7}$) (**Table 4**). The correlation between the whole blood MS and PGS in adults was -0.03 (*p*=0.182; **Supplementary Figure S4A**). In the model combining MS and PGS to predict birth weight, the PGS and blood MS in adults both significantly explained variation in birth weight (variance explained by MS in combined model: 0.39%, *p*=0.003; by PGS: 1.53%, *p*=$1.96x10^{-7}$ and MS+PGS: 1.92%) (**Table 5**, **Figure 1E**).

In children, the best performing score was based on 958 CpGs with a p-value lower than $1x10^{-7}$, explaining 0.14% of the variance (*p*=0.263) (**Table 3**). The PGS predicted birth weight in children (variance explained by PGS=1.39%, *p*=$9.67x10^{-5}$) (**Table 4**). MSs did not add predictive value to PGS in the combined model (**Table 5**, **Figure 1F**).

TABLE 3.
Results of the methylation score prediction of birth weight, BMI, prenatal maternal smoking and current and former smoking

Trait	Group – DNA methylation tissue	CpGs included in MS	*N* CpGs	β_{MS}	SE_{MS}	P_{MS}	MS R^2 (%)
Birth weight	Adults – Whole blood (*N*=2,040)	$p<10^{-1}$	72,570	0.044	0.029	0.127	0.192
		$p<10^{-5}$	2,274	0.055	0.029	0.057	0.304
		$p<10^{-7}$	963	0.049	0.027	0.074	0.238
		$p<10^{-1}$ pruned**	934	0.062	0.022	0.004	0.386
		$p<10^{-5}$ pruned	30	0.015	0.023	0.513	0.023
		$p<10^{-7}$ pruned	18	0.025	0.023	0.277	0.064
	Children – Buccal cells (*N*=1,070)	$p<10^{-1}$	72,205	-0.013	0.031	0.679	0.016
		$p<10^{-5}$	2,249	0.031	0.032	0.344	0.094
		$p<10^{-7}$ **	958	0.038	0.032	0.237	0.141
		$p<10^{-1}$ pruned	184	0.015	0.029	0.613	0.022
		$p<10^{-5}$ pruned	13	0.016	0.032	0.613	0.025
		$p<10^{-7}$ pruned	9	-0.005	0.033	0.882	0.002
BMI	Adults – Whole blood (*N*=2,410)	$p<10^{-1}$	55,653	0.134	0.025	6.98×10^{-08}*	1.786
		$p<10^{-5}$	1,067	0.261	0.026	5.07×10^{-24}*	6.822
		$p<10^{-7}$ **	412	0.277	0.026	9.79×10^{-27}*	7.673
		$p<10^{-1}$ pruned	671	0.124	0.021	3.14×10^{-09}*	1.538
		$p<10^{-5}$ pruned	13	0.220	0.023	1.70×10^{-22}*	4.827
		$p<10^{-7}$ pruned	6	0.206	0.023	7.50×10^{-20}*	4.258
	Children – Buccal cells (*N*=1,072)	$p<10^{-1}$	55,279	0.003	0.017	0.878	0.001
		$p<10^{-5}$	1,079	-0.006	0.017	0.733	0.003
		$p<10^{-7}$	422	-0.008	0.018	0.676	0.006
		$p<10^{-1}$ pruned**	183	0.021	0.019	0.276	0.042
		$p<10^{-5}$ pruned	13	0.005	0.019	0.772	0.003
		$p<10^{-7}$ pruned	6	0.007	0.018	0.720	0.004
Prenatal Maternal Smoking	Adults – Whole blood (*N*=720)	$p<10^{-1}$	76,531	0.761	0.168	5.83×10^{-06}*	8.512
		$p<10^{-5}$	1,581	0.946	0.140	1.51×10^{-11} *	15.115
		$p<10^{-7}$ **	607	1.009	0.136	1.28×10^{-13} *	17.277
		$p<10^{-1}$ pruned	962	0.580	0.136	2.01×10^{-05} *	7.076
		$p<10^{-5}$ pruned	33	0.661	0.123	7.12×10^{-08}*	6.820
		$p<10^{-7}$ pruned	16	0.620	0.119	1.94×10^{-07} *	5.963

Table 3. (continued)

Trait	Group – DNA methylation tissue	CpGs included in MS	*N* CpGs	β_{MS}	SE_{MS}	P_{MS}	MS R^2 (%)
Smoking Current Never	Children – Buccal cells (*N*=547)	$p<10^{-1}$	76,146	-0.256	0.158	0.105	0.930
		$p<10^{-5}$	1,571	-0.285	0.187	0.127	1.683
		$p<10^{-7}$ **	606	-0.304	0.183	0.096	2.223
		$p<10^{-1}$ pruned	187	-0.125	0.153	0.416	0.386
		$p<10^{-5}$ pruned	16	-0.096	0.164	0.559	0.149
		$p<10^{-7}$ pruned	9	-0.120	0.161	0.456	0.274
	Adults – Whole blood (*N*=1,914)	$p<10^{-1}$	98,972	1.420	0.095	$4.65x10^{-43*}$	20.974
		$p<10^{-5}$	11,433	1.923	0.104	$1.93x10^{-66*}$	40.086
		$p<10^{-7}$	6,938	1.999	0.105	$6.89x10^{-69*}$	44.449
		$p<10^{-1}$ pruned	913	1.170	0.070	$2.70x10^{-47*}$	27.626
		$p<10^{-5}$ pruned	37	2.194	0.108	$5.21x10^{-75*}$	56.237
		$p<10^{-7}$ pruned **	24	2.245	0.111	$1.25x10^{-73*}$	57.461
Smoking Former Never	Adults – Whole blood (*N*=1,938)	$p<10^{-1}$ from current versus never smoking EWAS	98,972	0.429	0.085	$6.63x10^{-07*}$	2.004
		$p<10^{-5}$ from current versus never smoking EWAS	11,433	0.689	0.086	$2.73x10^{-15*}$	5.127
		$p<10^{-7}$ from current versus never smoking EWAS	6,938	0.758	0.086	$3.37x10^{-18*}$	6.217
		FDR significant from Former vs Never Smoking EWAS	2,568	0.935	0.088	$3.56x10^{-24*}$	9.340
		$p<10^{-1}$ pruned from current versus never smoking EWAS	913	0.499	0.066	$3.38x10^{-13*}$	4.399
		$p<10^{-5}$ pruned from current versus never smoking EWAS	37	1.186	0.087	$6.64x10^{-36*}$	15.625
		$p<10^{-7}$ pruned from current versus never smoking EWAS**	24	1.226	0.088	$5.44x10^{-36*}$	16.316
		FDR significant from Former vs Never Smoking EWAS, pruned	2,330	0.692	0.071	$1.96x10^{-18*}$	7.569

β is the regression coefficient for each methylation score (MS) with standard error (SE) and *p*-value (*P).* MS R2 is the phenotypic variance explained by the MS.

* indicates *p*-values <0.0012 in adults and <0.0016 in children

** indicate methylation score with lowest *p*-value for a trait/exposure

TABLE 4.
Results of the polygenic score prediction of birth weight, BMI, prenatal maternal smoking and current and former smoking

Trait	Group	β_{PGS}	SE_{PGS}	P_{PGS}	PGS R^2 (%)
Birth weight	Adults (*N* = 2,040)	0.123	0.024	$1.56x10^{-07*}$	1.520
	Children (*N* = 1,070)	0.118	0.030	$9.67x10^{-05*}$	1.387
BMI	Adults (*N* = 2,410)	0.259	0.022	$3.43x10^{-32*}$	6.725
	Children (*N* = 1,072)	0.173	0.041	$2.20x10^{-05*}$	3.003
Prenatal Maternal Smoking	Adults (*N* = 720)	0.259	0.132	0.049	1.797
	Children (*N* = 547)	0.282	0.165	0.086	1.622
Smoking Current vs Never	Adults (*N* = 1,914)	0.330	0.056	$2.24x10^{-07*}$	2.794
Smoking Former vs Never	Adults (*N* = 1,938)	0.197	0.058	0.001	0.909

β is the regression coefficient for each polygenic score (PGS) with standard error (SE) and *p*-value *(P).* PGS R2 is the phenotypic variance explained by the PGS.
PGS to predict current versus never smoking and former versus never smoking were created based on GWAS on smoking initiation[17].
* indicates *p*-values <0.0012 in adults and <0.0016 in children.

BMI

Blood MSs for BMI were based on the EWAS by Wahl et al.[12] in blood DNA in adults. These account for a moderate proportion of the variance in BMI in adults (1.5–7.7%). The best performing score explained 7.7% of the variance in BMI ($p= 9.79x10^{-27}$) and was based on 412 Bonferroni significant CpG sites (**Table 3**, **Figure 1B**). The pruned MSs explained less variation in BMI (1.5–4.8% explained variance). The PGS for BMI explained 6.7% of the variance in adults ($p=3.43x10^{-6}$; **Table 4**). The correlation between whole blood MS and PGS was 0.1 in adults ($p=1.3x10^{-6}$; **Supplementary Figure S4B**). In a combined regression model in adults the MS and PGS contributed independently to the prediction of BMI (variance explained by MS in combined model: 6.4%, $p=5.46x10^{-23}$, by PGS: 5.5%, $p=7.02x10^{-28}$, and MS+PGS: 11.9%) (**Table 5**; **Figure 1E**).

In children, the BMI MSs based on buccal methylation data had a considerably lower predictive performance, and none of the scores significantly predicted BMI: the best score in children explained 0.04% of the variance ($p=0.276$), and was based on 183 pruned CpG sites with a p-value lower than $1x10^{-1}$. The PGS explained 3% of the BMI variance ($p=2.2x10^{-5}$). MSs did not outperform PGSs in the combined model (**Table 5**; **Figure 1F**). Furthermore, removal of children for whom information on BMI was collected more than 3 years before or after DNA collection did not lead to an increase in explained variance (**Supplementary Table S4**).

Prenatal maternal smoking

In adults, the MSs were based on the EWAS in cord blood from 4,994 newborns 19,51 and significantly predicted prenatal maternal smoking exposure (**Table 3**, **Figure 1C**). The score based on weights of 607 unpruned CpGs at p-value $<1x10^{-7}$ accounted for largest variance of 17.3% of prenatal maternal exposure ($p=1.28x10^{-13}$). The pruned MSs

performed worse (5.9-7% explained variance). The PGS for maternal smoking around birth did not significantly predict prenatal maternal smoking (variance explained by PGS 1.8%, p=0.05) (**Table 4**). The correlation between the best performing MS and PGS was 0.06 (p=0.003; **Supplementary Figure S4C**). The variance explained by MS and PGS in the combined model was slightly lower than predicted by MS alone (variance explained by MS in combined model: 16.9%, p=2.4x10^{-13}, by PGS: 0.84%, p=0.139 and by MS+PGS: 17.7%) (**Table 5**; **Figure 1E**). Maternal smoking scores in buccal methylation data from children, based on the same cord blood discovery EWAS, were not significantly predictive (**Table 5**; **Figure 1F**).

Smoking

The smoking MS in adults were based on the EWAS for never versus current smokers[18,51] and were strongly predictive for smoking status. The pruned MSs had a considerably better predictive performance (28-57.5% vs 21-44% explained variance). The best performing MSs was based on 24 pruned CpGs at p-value <1x10^{-7}, and explained 57.5% of variance for current smoking (p=1.25x10^{-73}) and 16.3% of the variance for former smoking (p=5.44x10^{-36}) (**Table 3**, **Figure 1D**). The PGS for smoking initiation explained 2.8% of the variance in current smoking (p=2.24x10^{-7}) and 0.9% of the variance of former smoking (p=0.001) (**Table 4**). The correlation between the PGS of smoking initiation and the best performing MS was 0.14 (p=7.71x10^{-13}; **Supplementary Figure S4D**). In the combined prediction model, the MSs outperformed PGSs in the prediction of smoking status, and the PGSs were no longer significant (variance explained by MS in combined model for current vs never smoking: 54.4%, p=1.3x10^{-70}, by PGS: 0.16%, p=0.042, and by MS+PGS: 54.6%), indicating that the PGS and MS for smoking do not independently add to prediction of this trait (**Table 5**, **Figure 1E**).

Discussion

We examined if a combined model that includes methylation scores (MS) and polygenic scores (PGS) captures more variance in body size, i.e., birth weight and BMI, and in two exposures, i.e., prenatal maternal smoking exposure and smoking in adulthood, in comparison to the PGS alone. Our results showed that MSs in adults, from blood DNA, predicted BMI, prenatal maternal smoking, and smoking status independent of PGSs, and outperformed PGSs for BMI, prenatal maternal smoking, and smoking status, but not for birth weight. In children, MSs from buccal-cell DNA did not show predictive value in children, but here the tissue in the discovery studies derived from EWASs of cord blood and whole blood DNA methylation profiles.

The most successful MS predictor in our study is for smoking. Blood DNA MS explained up to 57.5% of the variance in current smoking status and 16.3% of the variance in former smoking status. This was substantially better compared to the performance of PGS. Tobacco exposure, both prenatal and current, is a potential environmental exposure that modifies DNA methylation. Several previous studies reported successful application of blood DNA MS created based on weights from an independent discovery EWAS, as we did in the current study[11,23,26], based on calculation of indexes[27,28] or based

TABLE 5.
Results of the multi-omics prediction of birth weight, BMI, prenatal maternal smoking and current and former smoking

Trait	Group – DNA methylation tissue	Methylation score					Polygenic score				Total R^2 for MS+PGS (%)
		CpGs included in MS	β_{MS}	SE_{MS}	P_{MS}	MS R^2 (%)	β_{PGS}	SE_{PGS}	P_{PGS}	PGS R^2 (%)	
Birth weight	Adults – Whole Blood (*N* = 2,040)	$p<10^{-1}$ pruned	0.063	0.021	0.003	0.394	0.124	0.024	$1.96x10^{-07*}$	1.525	1.92
	Children – Buccal cells (*N* =1,070)	$p<10^{-7}$	0.043	0.032	0.172	0.186	0.110	0.030	$2.61x10^{-04*}$	1.219	1.41
BMI	Adults – Whole Blood (*N* = 2,410)	$p<10^{-7}$	0.253	0.026	$5.5x10^{-23*}$	6.40	0.235	0.022	$7.02x10^{-28*}$	5.54	11.95
	Children – Buccal cells (*N* =1072)	$p<10^{-1}$ pruned	0.017	0.019	0.347	0.030	0.157	0.042	$1.7x10^{-04*}$	2.451	2.48
Prenatal Maternal Smoking	Adults – Whole Blood (*N* =720)	$p<10^{-7}$	1.031	0.140	$2.4x10^{-13*}$	16.886	0.207	0.140	0.139	0.837	17.72
	Children – Buccal cells (*N* =547)	$p<10^{-7}$ pruned	-0.291	0.187	0.120	1.529	0.283	0.168	0.091	1.093	2.62
Smoking Current vs Never	Adults – Whole Blood (*N* =1,914)	*p*<10-7 pruned from current versus never smoking EWAS	2.251	0.113	$1.3x10^{-70*}$	54.41	0.165	0.080	0.042	0.16	54.57
Smoking Former vs Never	Adults – Whole Blood (*N* =1,938)	0.197	1.216	0.089	$1.5x10^{-34*}$	15.52	0.159	0.064	0.018	0.51	16.03

β is the regression coefficient for each term with standard error (SE) and *p*-value *(P)*. MS R2 is the phenotypic variance explained by the MS. PGS R^2 is the phenotypic variance explained by the PGS. The combined model included the PGS and the best performing MS. * indicates *p*-values <0.0012 in adults and <0.0016 in children.

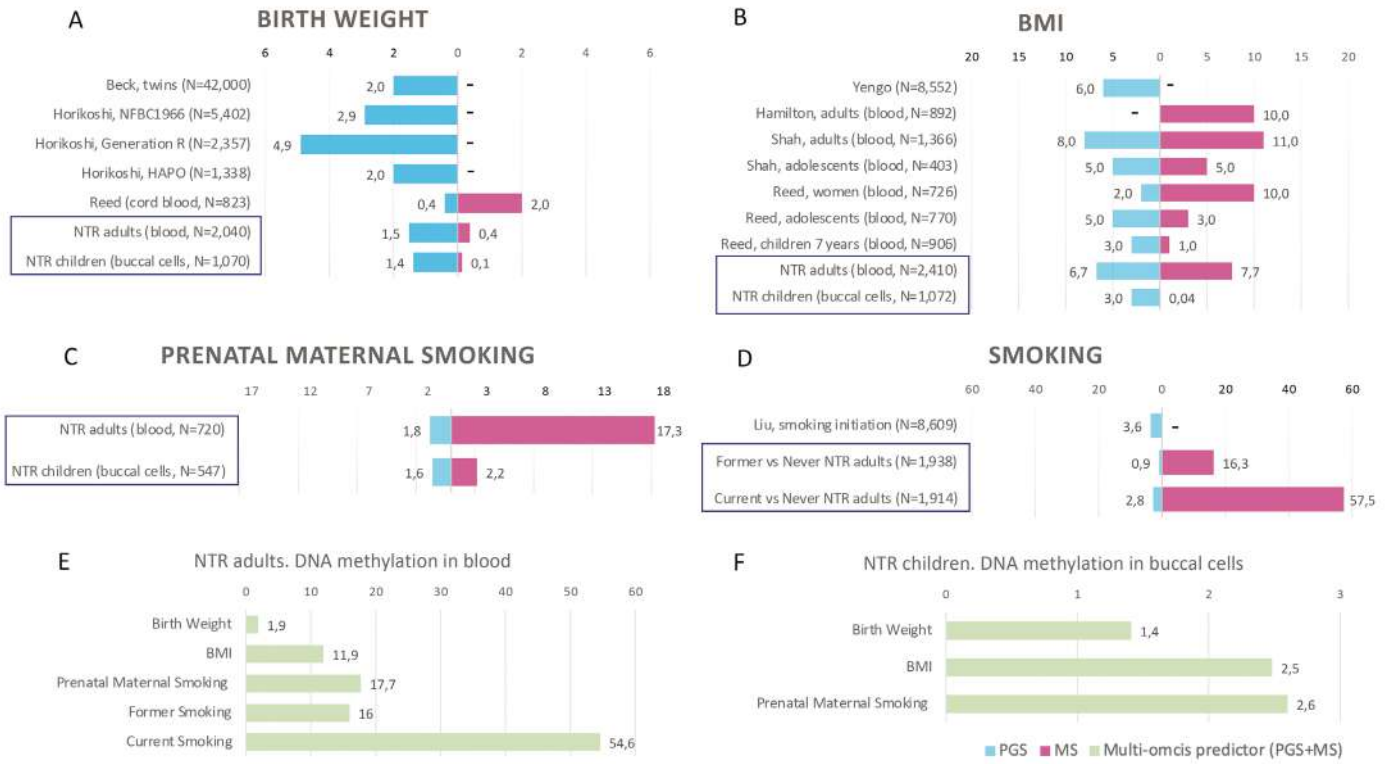


Figure 1. Prediction by methylation and polygenic scores in previous studies and NTR

Bars are the phenotypic variance explained by the score (R^2), x-axis shows R^2 in %. MS, methylation score. PGS, polygenic score. PGS+MSpolygenic and methylation scores in combined model (multi-omics predictor). "–" indicates that the score is not available in the study. Prediction by PGS and MS in NTR cohorts is indicated by blue frames in **A-D**. Multi-omics prediction in NTR is presented in **E-F**. Full references on previous studies in **A, B, D** can be found by first author in References. For more details on previous EWASs included in **A, B** see **Supplementary Table S1**.

on machine learning algorithms[25]. In line with previous studies, MSs performed better for predicting current vs never smoking than for former vs never smoking[25]. Most studies of smoking were done on blood DNA methylation. It has been suggested that buccal cell DNA methylation predictors should perform even better[25]. Currently, our participants with buccal cell DNA methylation data are too young (methylation data in buccal cells was available for children around 9 years old) to have initiated smoking.

The blood DNA MSs for prenatal maternal smoking, based on cord blood-derived weights from newborns, significantly explained 17.3% of variance in adults. Earlier reports demonstrated that maternal smoking during pregnancy is associated with alterations in offspring blood DNA methylation in newborns[19,55,56], children and adolescents[57,58], and adults several decades after exposure[23,59]. However, the effects of sustained maternal smoking during pregnancy fade away with time, and the predictive accuracy in blood samples from adults is much lower than the accuracy obtained with cord blood samples from newborns[56–58,60].

We showed that MSs perform better than PGSs in adults for both exposures capturing the effect of smoking, and add value to prediction in combined models. The effects of individual SNPs on behavioral traits such as smoking is small, hence, larger GWAS meta-analyses are required for smoking and maternal smoking to obtain better PGSs. In contrast to PGSs, which capture an individual's genetic predisposition for smoking behavior, MSs capture the effect of exposure to smoking on the methylome. Smoking as an exposure is strongly associated with DNA methylation, and EWAS meta-analyses of smoking and maternal smoking have identified very large numbers of CpGs associated with these traits, allowing for the calculation of fairly reliable MSs, i.e., the EWAS meta-analysis identified over 2,000 significant loci associated with smoking[18], while the currently largest GWAS detected 566 loci associated with smoking initiation[17].

The BMI MS derived in blood samples from adults explained up to 7.7% of the variation in BMI, thereby outperforming the PGS, which explained 6.7% of the variance. Both scores contributed independently to the prediction of BMI in the combined model (12% of explained variance). The performance of adult NTR MSs for BMI was in line with other studies that reported around 10-11% variance explained by MS only in adults (see **Figure 1B**) and larger variance explained by combined MS and PGS predictors[21,22]. In children with buccal methylation data, a considerably smaller proportion of 2.5% of variation in BMI was explained by the MS and PGS in combined predictor. The lower predictive performance of BMI MSs in children than in adults was also observed in other studies[20,22], and could be explained by increase of environmental contribution to the trait with age[61] as BMI tends to increase during most of adult life[62]. Shah et al.[22] reported that BMI MSs based on an EWAS in adults from the Lothian Birth Cohorts explained 4.9% of variation in adults from the Lifelines DEEP study, but did not account for any BMI variation in adolescents (mean age 14 years) from the Brisbane Systems Genetic Study. Reed et al.[20] observed 10% of BMI variance explained in women in comparison with 1% in children age 7 years and 3% in adolescents age 15 years by MSs calculated on the same set of CpGs from an EWAS in adults.

Birthweight MSs were not strongly predictive in our study, with 0.4% and 0.1% of explained variance in adults and children respectively, while PGS were significant with

1.5% and 1.4% of explained variance in adults and children respectively. The PGS for birth weight in NTR was in between the variance explained in previous studies (see **Figure 1A**): 0.4% in an ALSPAC cohort of 823 newborns[20], 2% in multi-cohort study of 42 thousand twins[53] and in Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study of 1,338 individuals14, 2.9% in the Northern Finland Birth Cohort (NFBC) of 5,402 individuals and 4.9% in Generation R cohort of 2,357 individuals[14]. The discovery EWAS of birth weight (N = 8,825 newborns) detected 914 Bonferroni significant CpGs in cord blood[15], suggesting that birth weight does have a large epigenetic signal in cord blood at birth. The cord-blood DNA MS base on weights from adult BMI EWAS accounted for 2% of variance in birth weight[20]. According to our knowledge, the performance of birthweight MS based on weights from newborn EWAS has not been previously examined. The low predictive accuracy of the MSs can be caused by the fact the birth weight scores in whole blood and buccal cells were based on an EWAS in cord blood of neonates. Across different tissues and ages, different CpG sites may be associated with birth weight. Another explanation is, that the association between birth weight and DNA methylation fades away with age. Küpers et al.[15] took the 914 significant neonatal blood CpG sites and examined their associations with birth weight in blood samples of adults. No CpG site reached Bonferroni significance in the adults. At present, there are no published EWASs of birth weight based on buccal DNA methylation and no large EWASs of birth weight on blood samples from adults.

The lower predictive accuracy of the buccal cell DNA MSs for all four phenotypes may have several reasons: (1) for some traits there is evidence that DNA methylation signatures increases with age, e.g., BMI, and thus can be not evident at age of 9 years old; (2) unreliability in the phenotype, e.g., prenatal maternal smoking reported participants on their mother's smoking behavior; (3) use of effect sizes from EWASs in cord blood and whole blood methylation data to calculate the scores in buccal cell DNA methylation data. The CpG sites that are predictive for trait/exposure in blood may not be the same CpG sites that are predictive in buccal cells, or the strength of the association may differ across tissues. This last explanation can be tested once more EWASs in buccal cells become available.

The best performing score for each trait will depend on the true number of CpGs associated with the trait and their effect size, the correlation among CpGs, and the power of the discovery EWAS analysis. If the discovery EWAS had full power to detect all CpGs associated with the trait, and there is no large heterogeneity in the effects across cohort, scores created within the same tissue as the discovery EWAS with CpGs based on the most stringent p-value threshold (i.e., $< 1x10^{-7}$) are expected to perform best. More likely, EWASs for these traits did not yet detect all CpGs that are truly associated and larger discovery samples are required to detect CpGs with smaller effects. Therefore, we also examined the performance of scores created based on more lenient p-value thresholds. More lenient p-value thresholds will potentially add more CpGs to the MS that are truly associated with the phenotype, but which did not yet reach epigenome-wide significance in the discovery EWAS meta-analysis, thereby improving the score. At the same time, inclusion of more CpGs that are not truly associated with the trait and less accurate weights at more lenient thresholds, add more noise to the MS. Pruning was performed to remove correlated CpGs that are redundant (and potentially add noise to

scores). The expectation is that if the set of CpGs associated with a trait is correlated (and especially if correlations are strong or abundant), pruning will improve performance of the MS. We found this to be the case, for instance, for blood DNA MS for smoking in adults. For simplicity, we compared two options that have been previously applied in the literature: (1) no pruning at all and (2) a correlation cut-off of 0.1 to select an approximate independent set of CpGs, but we note that the optimal correlation cut-off for pruning may also vary across traits. In adults, pruning reduced the performance of some scores, namely BMI and prenatal maternal smoking, while it improved the performance for birth weight and smoking. Sophisticated methods for MS calculation that model the exact correlation structure between CpGs, as are available for PGS[52], are yet to be developed. In our study, we have selected the best performing score for each trait based on the currently available largest EWAS. With larger discovery EWASs, the optimal selection approach for CpGs is also expected to change.

Birth weight and BMI are physical characteristics, whereas prenatal maternal smoking and own smoking are commonly labeled as exposures and behavioral traits. Birth weight is the least heritable of these traits, while mother's behavior in prenatal maternal smoking consists of an exposure whose genetic contribution is genetically transmittable to offspring, who inherit 50% of mothers' genes. All four complex traits are influenced by genetic variants and environmental factors, although some have argued that behavioral traits are more distal and less directly under biological control than physical traits. Polygenic signals from PGS and MS are composites of signals from different sources that are a result of different combinations of underlying biological processes. Notwithstanding the gap in our understanding about biological processes between the polygenic signals and phenotypes and exposure outcomes, the hypothesis-free approaches from GWAS and EWAS allow for construction of polygenic and methylation scores that have certain predictive accuracy, as demonstrated in research and that have potential for clinical use[10,63].

The pathways between genome and complex physical and behavioral traits may pass over many different cascades of biological processes in interplay and interaction with environmental factors. PGS and MS can capture different sources of information, from GWASs and EWASs. PGS will capture only genetic vulnerability for a trait, while MSs may capture, in addition to genetic influences on the trait, environmental and stochastic influences and the effect of the trait on the MS.

Pathway analyses indicate that protein products of genes within birthweight-associated regions in GWAS are enriched for diverse processes including insulin signaling, glucose homeostasis, glycogen biosynthesis and chromatin remodeling[14]. Birthweight-associates CpGs are among sites that have previously been linked to prenatal maternal smoking and mother's BMI before pregnancy[15]. Genes annotated to BMI-associated SNPs are mostly enriched among genes involved in neurogenesis and more generally in the development of the central nervous system[16]. Cell type-specific gene expression analysis identified enrichment of brain cell types in BMI[64]. These findings suggest that BMI could be considered as a behavioral trait and not only metabolic one. Genes annotated to BMI-associated CpGs play role in adipose tissue biology, insulin resistance, inflammation, as well as metabolic, cardiovascular, respiratory and neoplastic disease[12].

Smoking-initiation-associated genes are involved in dopaminergic and glutamatergic neurotransmission among several regions in the central nervous system related to addictive behavior[17]. Many CpGs overlap in both in newborns exposed to prenatal maternal smoking and smoking adults (including cg05575921 *(AHRR)* indicative to smoking exposure in many studies) are implicated in numerous neurological pathways, embryogenesis, and various developmental pathways[51]. Unique pathways observed in newborns include xenobiotic-related pathways, cytochrome P450 and uridine-glucoronosyltansferases involved in metabolism of nicotine and other compounds of tobacco smoke[51], and pathways associated to susceptibility to orofacial clefts[19]. Unique pathways observed in adults EWAS are enriched for variants associated in GWAS with smoking-related disease, including osteoporosis, colorectal cancers, and chronic obstructive pulmonary disease.

The limitations of our study relate to measurement reliability and the missing data for some phenotypes, which reduced the study power. The largest number of missing data was for prenatal maternal smoking in adults (47%, *N*=1,133). Women may underreport smoking during pregnancy[24], although in NTR the prevalence of maternal smoking during pregnancy was 19.5 % for the mothers of young twins, which is in line with the prevalence reported in the general Dutch population[65]. Birth weight data were missing for many participants who joined the NTR as adults (17%, *N*=404). In children, there were varying time differences between DNA methylation and BMI measurement. Our sensitivity analysis showed that MS prediction accuracy was not affected by this difference. In the prediction models, we adjusted for age, sex, and cellular composition of samples, hence the predictive performance of MSs reported in this paper is over and above the effects of age, sex, and cellular composition. We recognize that the MSs, and their ability to predict the phenotypes that we study are likely to be impacted by other factors, such as gestational age for birth weight and prenatal maternal smoking, BMI and amount of cigarettes in smoking exposure and vice versa smoking in BMI. Further explorations of potential confounders and mediators will be valuable.

The combination of PGS and MS is a tool to address research questions, such as mediation by DNA methylation of the effect of certain exposures on a trait of interest, where a score based on multiple CpGs may increase the power of such studies compared to a single CpG site.

Lifestyle variables, such as smoking behavior, are often assessed in epidemiologic studies by interviews or questionnaires, and individuals may hide their smoking status or adults may not know if their mother smoked during pregnancy. In such cases, the use of epigenetic profiles can serve as biomarkers and be applied an alternative of survey data[11]. Further, the MSs also have potential to be used in risk stratification and disease risk prediction. For example, BMI MSs were shown to predict type 2 diabetes beyond traditional risk factors including BMI and waist–hip ratio[12].

In conclusion, this study illustrates the value of combining PGS with MS for complex trait and exposure prediction. The results of our study provide new insights into the predictive performance of PGS and MS for different traits, across different tissues and ages. Because we analyzed buccal data in NTR children and blood data in NTR adults, the

current study could not distinguish between age and tissue as cause for the differences in predictive performance of the scores in the two groups. To make a better distinction between differences caused by age or tissue type, future studies that can create PGS and MS based on both blood and buccal data in children and adults are warranted and ideally both tissues are available for the same individuals. Furthermore, the predictive performance of MSs in blood and buccal methylation data may improve if MSs will be created based on EWA studies performed in the same type of tissue collected at the same age with larger sample size, with other approaches rather than weighted score approach (e.g., machine learning), and the prediction of traits may be further improved by adding information from additional omics levels. Future follow-up studies should investigate relationships between the DNA sequence and DNA methylation in complex traits and exposure outcomes.

References

*Discovery studies used to calculate PGS and MS in this article are indicated with *.*

1. Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. *Nature reviews. Genetics* **3**, 872–882 (2002).
2. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature genetics* **47**, 702–709 (2015).
3. Van Dongen, J., Slagboom, P. E., Draisma, H. H. M., Martin, N. G. & Boomsma, D. I. The continuing value of twin studies in the omics era. *Nature Reviews Genetics* **13**, 640–653 (2012).
4. Yong, S. Y., Raben, T. G., Lello, L. & Hsu, S. D. H. Genetic architecture of complex traits and disease risk predictors. *Scientific reports* **10**, 12055 (2020).
5. Harden, K. P. & Koellinger, P. D. Using genetics for social science. *Nature human behaviour* **4**, 567–576 (2020).
6. Smith-Woolley, E., Selzam, S. & Plomin, R. Polygenic score for educational attainment captures DNA variants shared between personality traits and educational achievement. *Journal of personality and social psychology* **117**, 1145–1163 (2019).
7. Toulopoulou, T. *et al.* Polygenic risk score increases schizophrenia liability through cognition-relevant pathways. *Brain* **142**, 471–485 (2018).
8. Khera, A. *et al.* Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. *Cell* **177**, 587-596.e9 (2019).
9. Mavaddat, N. *et al.* Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *The American Journal of Human Genetics* **104**, 21–34 (2019).
10. Hüls, A. & Czamara, D. Methodological challenges in constructing DNA methylation risk scores. *Epigenetics* **15**, 1–11 (2020).
11. Sugden, K. *et al.* Establishing a generalized polyepigenetic biomarker for tobacco smoking. *Translational psychiatry* **9**, 92 (2019).
12. Wahl, S. *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* **541**, 81–86 (2017).*
13. Rask-Andersen, M., Karlsson, T., Ek, W. E. & Johansson, Å. Gene-environment interaction study for BMI reveals interactions between genetic factors and physical activity, alcohol consumption and socioeconomic status. *PLOS Genetics* **13**, e1006977 (2017).
14. Horikoshi, M. *et al.* Genome-wide associations for birth weight and correlations with adult disease. *Nature* **538**, 248–252 (2016).
15. Küpers, L. K. *et al.* Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. **10**, 1893 (2019). *
16. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Human molecular genetics* **27**, 3641–3649 (2018). *
17. Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature genetics* **51**, 237–244 (2019).*
18. Joehanes, R. *et al.* Epigenetic Signatures of Cigarette Smoking. Circulation. *Cardiovascular genetics* **9**, 436–447 (2016). *
19. Joubert, B. R. *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *The American Journal of Human Genetics* **98**, 680–696 (2016).

20. Reed, Z. E., Suderman, M. J., Relton, C. L., Davis, O. S. P. & Hemani, G. The association of DNA methylation with body mass index: distinguishing between predictors and biomarkers. *Clinical epigenetics* **12**, 50 (2020).

21. Hamilton, O. K. L. *et al.* An epigenetic score for BMI based on DNA methylation correlates with poor physical health and major disease in the Lothian Birth Cohort. *International Journal of Obesity* **43**, 1795–1802 (2019).

22. Shah, S. *et al.* Improving Phenotypic Prediction by Combining Genetic and Epigenetic Associations. *The American Journal of Human Genetics* **97**, 75–85 (2015).

23. Richmond, R. C., Suderman, M., Langdon, R., Relton, C. L. & Davey Smith, G. DNA methylation as a marker for prenatal smoke exposure in adults. *International journal of epidemiology* **47**, 1120–1130 (2018).

24. Reese, S. E. *et al.* DNA Methylation Score as a Biomarker in Newborns for Sustained Maternal Smoking during Pregnancy. **125**, 760–766 (2017).

25. Bollepalli, S., Korhonen, T., Kaprio, J., Anders, S. & Ollikainen, M. EpiSmokEr: a robust classifier to determine smoking status from DNA methylation data. *Epigenomics* **11**, 1469–1486 (2019).

26. Elliott, H. R. *et al.* Differences in smoking associated DNA methylation patterns in South Asians and Europeans. *Clinical epigenetics* **6**, 4 (2014).

27. Nwanaji-Enwerem, J. C. *et al.* Relationships of Long-Term Smoking and Moist Snuff Consumption With a DNA Methylation Age Relevant Smoking Index: An Analysis in Buccal Cells. *Nicotine & tobacco research: official journal of the Society for Research on Nicotine and Tobacco* **21**, 1267–1273 (2019).

28. Zhang, Y. *et al.* Smoking-Associated DNA Methylation Biomarkers and Their Predictive Value for All-Cause and Cardiovascular Mortality. *Environmental Health Perspectives* **124**, 67–74 (2016).

29. Mendelson, M. M. *et al.* Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. *PLoS medicine* **14**, e1002215 (2017).

30. Gao, X., Zhang, Y., Breitling, L. P. & Brenner, H. Relationship of tobacco smoking and smoking-related DNA methylation with epigenetic age acceleration. *Oncotarget* **7**, 46878–46889 (2016).

31. Nwanaji-Enwerem, J. C. & Colicino, E. DNA Methylation–Based Biomarkers of Environmental Exposures for Human Population Studies. *Current Environmental Health Reports* **7(2)**, 121–128 (2020)

32. Willemsen, G. *et al.* The Netherlands Twin Register Biobank: A Resource for Genetic Epidemiological Studies. *Twin Research and Human Genetics* **13**, 231–245 (2010).

33. Boomsma, D. I. Aggression in children: unravelling the interplay of genes and environment through (epi)genetics and metabolomics. *Journal of Pediatric and Neonatal Individualized Medicine (JPNIM)* **4**, e040251–e040251 (2015).

34. Bartels, M. *et al.* Childhood aggression and the co-occurrence of behavioural and emotional problems: results across ages 3–16 years from multiple raters in six cohorts in the EU-ACTION project. *European Child & Adolescent Psychiatry* **27**, 1105–1121 (2018).

35. Hagenbeek, F. A. *et al.* Urinary Amine and Organic Acid Metabolites Evaluated as Markers for Childhood Aggression: The ACTION Biomarker Study. *Frontiers in Psychiatry* **11**, (2020).

36. Van Dongen, J. *et al.* DNA methylation signatures of aggression and closely related constructs: A meta-analysis of epigenome-wide studies across the lifespan. *Molecular psychiatry* **26**, 2148–2162 (2021).

37. Van Dongen, J. *et al.* Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nature Communications* **7**, 11115 (2016).

38. Bonder, M. J. *et al.* Disease variants alter transcription factor levels and methylation of their binding sites. *Nature genetics* **49**, 131–138 (2017).

39. Van Iterson, M. *et al.* MethylAid: visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics (Oxford, England)* **30**, 3435–3437 (2014).

40. Chen, Y. *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **8**, 203–9 (2013).

41. Francioli, L. C. *et al.* Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nature genetics* **46**, 818–825 (2014).

42. Fortin, J.-P. *et al.* Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome biology* **15**, 503 (2014).

43. Sinke, L., van Iterson, M., Cats, D., Slieker, R. & Heijmans, B. DNAmArray: Streamlined workflow for the quality control, normalization, and analysis of Illumina methylation array data. (2019) doi:10.5281/ZENODO.3355292.

44. Ligthart, L. *et al.* The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-Family Designs. *Twin Research and Human Genetics* **22**, 623–636 (2019).

45. Van Dongen, J. *et al.* Epigenome-Wide Association Study of Aggressive Behavior. *Twin Research and Human Genetics* **18**, 686–698 (2015).

46. Pidsley, R. *et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biology 17, 208 (2016).

47. Van Beijsterveldt, C. E. M. *et al.* The Young Netherlands Twin Register (YNTR): Longitudinal Twin and Family Studies in Over 70,000 Children. *Twin Research and Human Genetics* **16**, 252–267 (2013).

48. Dolan, C. *et al.* Testing Causal Effects of Maternal Smoking During Pregnancy on Offspring's Externalizing and Internalizing Behavior. *Behavior genetics* **46**, (2015).

49. Zheng, S. C. *et al.* A novel cell-type deconvolution algorithm reveals substantial contamination by immune cells in saliva, buccal and cervix. *Epigenomics* **10**, 925–940 (2018).

50. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS genetics* **2**, e190 (2006).

51. Sikdar, S. *et al.* Comparison of smoking-related DNA methylation between newborns from prenatal exposure and adults from personal smoking. *Epigenomics* **11**, 1487–1500 (2019).*

52. Vilhjálmsson, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *American journal of human genetics* **97**, 576–92 (2015).

53. Beck, J. J. *et al.* Genetic Meta-Analysis of Twin Birth Weight Shows High Genetic Correlation with Singleton Birth Weight. *Human Molecular Genetics* (2021) doi:10.1093/hmg/ddab121.

54. Lee, S. H., Goddard, M. E., Wray, N. R. & Visscher, P. M. A better coefficient of determination for genetic profile analysis. *Genetic epidemiology* **36**, 214–224 (2012).

55. Markunas, C. A. *et al.* Identification of DNA Methylation Changes in Newborns Related to Maternal Smoking during Pregnancy. *Environmental Health Perspectives* **122**, 1147–1153 (2014).

56. Küpers, L. K. *et al.* DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *International journal of epidemiology* **44**, 1224–1237 (2015).

57. Richmond, R. C. *et al.* Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Human Molecular Genetics* **24**, 2201–2217 (2015).

58. Lee, K. W. K. *et al.* Prenatal Exposure to Maternal Cigarette Smoking and DNA Methylation: Epigenome-Wide Association in a Discovery Sample of Adolescents and Replication in an Independent Cohort at Birth through 17 Years of Age. *Environmental Health Perspectives* **123**, 193–199 (2015).

59. Wiklund, P. *et al.* DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. *Clinical epigenetics* **11**, 97 (2019).

60. Ladd-Acosta, C. *et al.* Presence of an epigenetic signature of prenatal cigarette smoke exposure in childhood. *Environmental research* **144**, 139–148 (2016).

61. Elks, C. E. *et al.* Variability in the heritability of body mass index: a systematic review and meta-regression. *Frontiers in endocrinology* **3**, 29 (2012).

62. Villareal, D. T., Apovian, C. M., Kushner, R. F. & Klein, S. Obesity in older adults: technical review and position statement of the American Society for Nutrition and NAASO, The Obesity Society. *The American Journal of Clinical Nutrition* **82**, 923–934 (2005).

63. Abdellaoui, A. & Verweij, K. J. H. Dissecting polygenic signals from genome-wide association studies on human behaviour. *Nature Human Behaviour* 1–9 (2021).

64. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature genetics* **47**, 1228–1235 (2015).

65. Dolan, C. V *et al.* Testing Causal Effects of Maternal Smoking During Pregnancy on Offspring's Externalizing and Internalizing Behavior. *Behavior genetics* **46**, 378–388 (2016).

Supplements

The supplementary materials are available online at

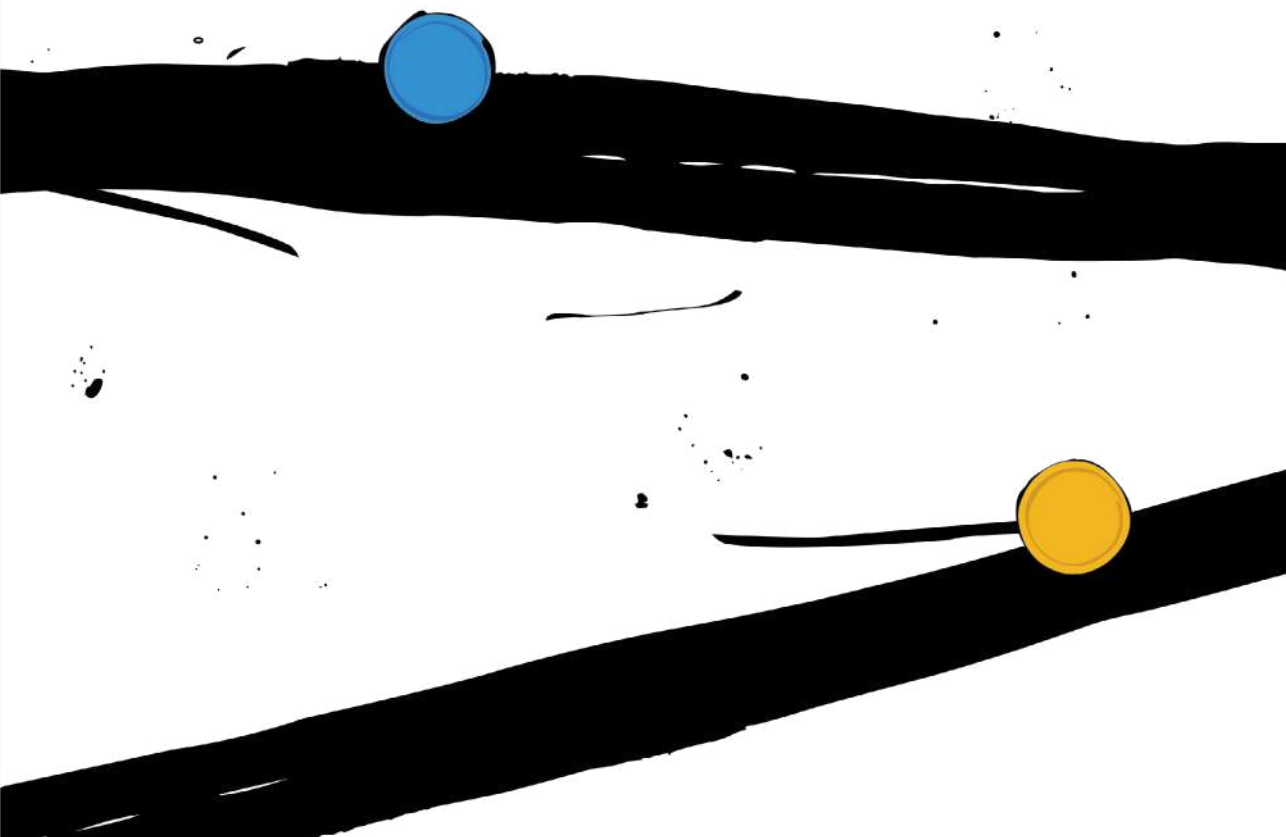
https://www.frontiersin.org/articles/10.3389/fpsyt.2021.688464/full#supplementary-material

PART III

Twin design for biomedical research: methodological issues

Chapter 9

Establishing a twin register: an invaluable resource for (behavior) genetic, epidemiological, biomarker and “omics” studies



Abstract

Twin registers are wonderful research resources for applications in epidemiology, molecular genetics and other areas of research. New registers continue to be launched all over the world as researchers from different disciplines recognize their potential. In this chapter, we discuss multiple aspects that need to be considered when initiating a register. This encompasses aspects related to the strategic planning and key elements of research designs, promotion and management of a twin register, including recruitment and retaining of twins and family members, phenotyping, database organization and collaborations between registers. Information on questions unique to twin registers and twin-biobanks, such as the assessment of zygosity by genotype arrays, the design of (biomarker) studies involving related participants and the analyses of clustered data, is presented. Altogether, we provide a number of basic guidelines and recommendations for reflection when planning a twin register.

Published as: Odintsova, V. V., Willemsen, G., Dolan, C. V., Hottenga, J. J., Martin, N. G., Slagboom, P. E., Ordoñana, J. R., & Boomsma, D. I. (2018). Establishing a twin register: an invaluable resource for (behavior) genetic, epidemiological, biomarker and "omics" studies. *Twin Research and Human Genetics,* 21(3): 239-252.

Introduction

Since the earliest times, it has been intuited that twins provide a window into the human condition, with numerous references in philosophical and literary texts, notably by Augustine of Hippo (400AD), Shakespeare, and others. The formalization of this intuition towards scientific study may best be attributed to Francis Galton in the late 19th century[1], although at that time it was not yet established that there were two discrete types of twins. Researchers in the early 1900s[2,3] pointed to the scientific value of twin studies[4], but in 1919 Sir Ronald Fisher still advocated that there was only one type of twins, stating that "...twins ordinarily share the hereditary nature of one gamete but not of the other"[2].

The first half of the 20th century saw a slow but steady development of twin research through the pioneering work of researchers as Poll[5] (1914), Merriman[6] (1924), Siemens[7] (1924), and Holzinger[8] (1929) (see Mayo[9] for a detailed account of this development). In Russia, twin studies were initiated as early as in 1900, with the first focus on psychological disturbances (see **Box 1**). Many twin studies were then undertaken and were essential in understanding the etiology of disorders. In France in the 1950's, Lejeune et al.[10], for example, became puzzled by the high concordance seen in Down syndrome in identical twins in comparison to the extremely low concordance in non-identical twins. This concordance pattern was inconsistent with single gene inheritance and was one of the observations which led to the discovery of trisomy-21 as the cause of Down syndrome. The history of twin studies, including the basic methodological insights and developments, has been described in multiple papers[4,9,11,12]. In the classical twin design, which includes mono- and dizygotic (MZ and DZ) twin pairs reared together, the resemblance for one or more human traits is compared between MZ and DZ twins to obtain estimates of uni and multivariate heritability. A larger resemblance in MZ twins is consistent with genetic influences on the trait under study. In the classical twin design, the statistical power is largest for detecting additive genetic influences (A). The two other variance components that contribute to resemblance of relatives, common or shared environmental variance (C), and dominance or non-additive genetic variance (D) require larger samples or the inclusion of additional family members to achieve reasonable power. Explorations of power of the classical twin study, first by simulation and later by direct analysis, showed that many thousands of pairs would be needed to separate these sources of variance[13,14]. This became the justification for the founding of large twin registries in a number of European countries, in the Scandinavian countries, the Netherlands, the United Kingdom, as well as in the United States and Australia. Many other countries in different parts of the world have followed suit. Compilations of twin registers across the world have been carried out periodically and published in the journal *Twin Research and Human Genetics* (2002, 2006, and 2013) and elsewhere[15]. Many of these twin registries are longitudinal, population based, and sufficiently large for epidemiological studies.

Twin registries have been a resource for thousands of studies, estimating the relative impact of genetic and environmental factors on trait variation across a wide range of biomedical and social science disciplines[16,17]. Their potential, however, for disentangling the role of genetics in human traits goes much further, through different designs and investigative methods, from genetic epidemiology to molecular approaches[18–21]. In many of the large registers, data collection is undertaken by mailed questionnaire or by telephone

Box 1: Twin Research in Russian Science
Reviews on the history of twin research tend to focus on developments in Western European countries and the USA. However, early references to twin studies are also to be found in Russian science. Psychiatrists Sergey Sukhanov and Tihon Yudin studied the similarity of psychosis in twins from 1900[75], and several small twin studies of morphological, physiological and psychological characteristics were conducted from 1900 to 1929[76]. The Russian Medical and Biological Institute, which was created in 1929 and continued as the Medical and Genetic Institute from 1935 onwards, conducted systematic and large-scale twin research where more than 700 twin pairs were studied. The research was conducted by medical doctors, psychologists, and pedagogues under the guidance of Solomon Levit. A special kindergarten for twins was created in the institute, in which motor functions, different forms of memory, level of psychic development, attention, and intellect features were studied. The method of twin-control design was used to study effectiveness of pedagogic, medical, and psychological interventions[77]. Unfortunately, these studies were limited and prohibited at the end of 1930s, and were restarted in Russia only in 1970ies in laboratory of genetic psychophysiology created by Irina Ravich-Shcherbo[78,79].

interview and more recently, by online survey, and record linkage. Clinical twin studies were of necessity often smaller, requiring twins to visit research or medical facilities, or researchers visiting twins at home. In anticipating of the age of molecular genetics, many registers started DNA collections in the early 1980s, and these samples have become a highly valuable and much used resource for zygosity assessment, genetic linkage and association studies. Collecting multiple sources of biological material has enabled twin studies of epigenetics, transcriptomics, metabolomics, proteomics, microbiome, and a wealth of biomarkers.

This chapter aims to discuss several aspects in establishing a twin register. We will attempt to cover what is important and why, and how to develop such a scientific resource. In this endeavor, we will also refer to when to start and who should participate. Our aim is not to offer a checklist or a complete step-by-step technical guide, but rather to discuss the issues that, in our experience, should be addressed at the launch and in the management of a twin register.

The First Steps

The first question that should be addressed is: “Do we really need to start a new twin register?”. Establishing a twin register is a huge and long-lasting effort and, although it pays back in the long run, it is costly, both in terms of economic and personal investment. Hence, research objectives must be clearly defined. Working with twins has many compensations that go even beyond the classical methodological advantages (see **Box 2**), and establishing a register may appear to be the best choice. However, it may not be the only option. Already established twin registries with data, or willing to collect new information, may be open to collaboration. In fact, nearly always a twin researcher can be found with an interest in collaboration and in replication of results. Using this option will not only result in economy of effort, avoiding duplication, but the proposed project may also benefit from the experience of other twin researchers.

Still, there may be many good reasons to start a twin register, for instance, in specific countries or populations. In that case, some other focal questions arise, starting with

the question of initial funding. This is, of course, a relevant question, and the available options will depend on many, often local, factors. When applying for initial funds, it may be practical to adapt the objectives to limited resources by focusing on a specific research topic rather than putting in a more general appeal to establish a research infrastructure. It is often wise to refine the research agenda in order to meet two complementary objectives: obtaining meaningful results in the short-term and looking for synergy with other research groups. Planning a long and complex research question will delay results that will be needed, since future funding will probably depend on early success. Hence, it is important to select a main phenotype to study based on its originality, the interest in the question within the scientific community, the uniqueness of the sample, or the available resources for data collection (see **Box 3**). Complementarily, it is important to seek out researchers not only within other twin registers around the world, but also outside the twin community, who have an interest in the selected phenotype and/or have relevant data. Collaboration with experienced researchers in the field is of value for a new project, while researchers from different disciplines may be interested in the possibilities that collaboration with a twin cohort offers. Identifying possible topics of common interest to the newly starting twin register and existing groups, which can contribute specific knowledge or techniques, may open new perspectives and facilitate trade-offs.

Strategic Planning

A twin register ideally is a longitudinal resource and, therefore, the first steps should be considered as the basis of a long-term effort. Decisions made during the first steps should facilitate the strategic planning of the register as a long-lasting organization. This involves setting the main goals and selecting the activities, in accordance to the available resources, which need to be undertaken to achieve the established objectives. Here, we first discuss human resources and adaptability to changing conditions. Human resources are, obviously, a core element of a twin register. A group of highly motivated and coordinated researchers is needed to start and develop a twin register. Thus, the question of identifying who may provide valuable help and be willing to participate

Box 2: Do It with Twins

In 1982, David Lykken listed in his presidential address to the Society for Psychophysiological Research[80] several compelling reasons for doing research with twins, that are 'in addition to' the genetic analyses that the classical twin design allows:

- Twins are plentiful and easily recruited as experimental subjects.
- Twins are probably more representative of the general population than any other group.
- This representativeness is even more true of the families of twins.
- Twin data are invaluable to explore issues measurement. Any measure that shows high within-pair correlation among MZ twins deserves to be treated with respect.
- The method of co-twin control provides enhanced experimental power. Using one twin from each pair for the experimental group and the other for the control group provides a test of one's hypothesis that is as powerful as an experiment employing twice as many pairs of singletons.
- If one treats one's subjects properly, and keeps in touch, then it will be possible to bring many of them back repeatedly over the years to participate in additional experiments. This is useful not only for longitudinal research but as a method of enhancing each subsequent experiment with the information previously gathered on these same individuals.

in the endeavor becomes crucial. Two different kinds of human resources should be contemplated: established researchers, from inside and outside the twin research community, who can contribute with expertise, advice, and logistics in their respective fields; and researchers or support personnel, who will be in charge of developing and maintaining the register. While the former are important in providing support and visibility, the latter are essential, since they will take care of the multiple tasks involved in the daily running of the register, i.e., from planning and conducting data collection or updating contacts, to analyzing data or writing papers. Therefore, human resources management (including selection, training, and career development) with the objective of forming a reliable and enduring core group is paramount if the register is to go ahead. Flexibility, adaptability, and keeping an eye on opportunities are also relevant issues. In a changing environment, where critical aspects such as funding or collaborations may change constantly, it seems wise to contemplate different horizons and be able to quickly adapt research objectives to different scenarios. This implies too the capacity to keep going with limited resources while being prepared for incoming opportunities. Focusing only on long-term and complex research objectives may represent a handicap for register development in case of funding shortage or operational obstacles. Keeping in mind and planning parallel sets of objectives adapted to different conditions may help to overcome temporary difficulties.

Basic Elements

There are several key elements that are at the core of the development of a twin register and that will determine its endurance and scientific success.

Recruitment Methods

One of the foremost questions of every researcher willing to start a twin register relates to which are the best practices for optimum recruitment and retention methods. There is not one clear answer and there may be as many methods as established registers. Recruitment strategies depend on a research protocol that can specify, for example, age at recruitment, inclusion and exclusion criteria, recruitment group (e.g., parents of young

Box 3: How Many Twins?

For most of the 20th century, until about 1970, there were only vague notions of how big twin studies needed to be to provide useful estimates of the degree of genetic influence (heritability), and many of the early studies, including small numbers, gave highly inconsistent results when complex human traits were analyzed. In retrospect, we can see this was mainly due to studies being underpowered, although inaccurate zygosity diagnosis also played a role. “Is there a genetic contribution to scholastic performance?” was the motivating question behind the first Australian twin study on school examination results from 1967[81,82]. While the study of 150 twin pairs was fairly large by the standards of that time, it soon became apparent that it was far too small to reliably estimate all genetic and environmental sources of variation, specifically the separation of additive genetic (A) and common or shared environmental variance (C). Multiple analytical and simulation studies now provide detailed tables with the required numbers of twin pairs for continuous and categorical traits, often distinguishing between uni and multivariate designs[13,14,83].

twins, adolescent, and adult twins), and possibilities of the research team, which may be affiliated with an academic institute or a medical infrastructure.

Table 1 summarizes, in a non-exhaustive manner, some of the possibilities for recruitment of participants. They can be divided into four major groups: (1) existing databases managed by public (e.g., city council, educational or health systems) or private (e.g., hospitals or insurance companies) stakeholders; (2) institutions or organizations that have access to twins; (3) participants recruited through media, advertisement, and social events; (4) word-of-mouth and recruitment through enrolled participants of register. There are many ways to find and enlist twins and, within these categories, researchers should be creative in finding ways to invite participants to a register. Different citizen registers or records can provide information about twins (e.g., birth or military records). In some countries, sampling twin pairs is based on computerized population registers, either from direct information on multiple births or from applying algorithms based on sharing of date of birth, family name at birth, place of birth, and so on. A request to provide addresses of persons born from the same mother with an identical date of birth can be done by municipalities. In all cases, 'real' twins have to be distinguished from a larger subset of 'possible' or 'administrative' twins, as sharing the name of the mother and date of birth might occur by chance[22]. Next, parents of twins or twins need to be contacted, with an invitation to participate in the register. Population samples can also be obtained through collaboration with hospitals and schools. Records can be available at maternity hospitals, which may give an opportunity of direct recruitment of study participants. The recruitment through schools gives possibilities to obtain information on school achievements from teachers. Many registries collaborate with twins or parents of twins associations.

Other twin collections are gathered independently of centralized records or institutions and may depend more on the motivation of the twins or their parents. Recruitment through advertising has been used, as well as through mass media articles on twins and twin research in which information on major achievements is combined with continuing studies and contact information. Such approaches can be effective, and the possible effects of bias in non-randomly ascertained samples can be dealt with by statistical methods[23,24]. Twin pairs can be registered via completion of a registration form online by either the twins or their parents if they are under the age of legal consent. Other avenues of recruitment include offering of booklets to parents who expect twins. The exposure on twin research findings on general media also attracts new participants. Some registries organize social events (e.g., twin festivals, a range of exhibitions about twins, including photos and pictures). Common meetings of enrolled and new participants can benefit in the realization of a register and contribute to the strengthening of the role and value awareness in participants. A useful practice can be when participants give presentations about their own experience during meetings or on social media or web-sites.

Informed consent

Twin registers are set up with the aim of conducting multiples studies across a long period of time and generally collect a wide variety of data in their participants. While participants upon registration may agree at the start of the study to the general aims provided in information brochures and will consent to be approached in years to come,

TABLE 1.
Recruitment Methods of Twin Registries

Using existing databases with information on twins managed publicly or privately
Previous twin studies
Population registries
Birth records
Immunization registries
Different patient registries
Voter records
Military records
Collaboration with institutions and organizations
State public health resources (e.g., healthcare departments)
Hospitals, maternal hospitals, and outpatient clinics
Insurance companies
Schools
Orphanages and adoption agencies
Multiple birth associations
Twin clubs and associations
Recruitment through media and social events
Media, newspapers, TV, and radio
Advertisement
Information brochures
Website
Social Media
Scientific and social events (e.g., twin festivals and annual gatherings)
Word-of-mouth and through enrolled participants in register

Note: The table does not attempt to be exhaustive.

the initial consent will not cover all the data to be collected in the future. Participants should be kept informed of the ways in which their data are used and be provided with the option of withdrawal at any stage of the research. Researchers need to establish how they will meet the participants' rights to know and to withdraw. Although the way this is laid down in law will not be the same across all countries in the world, it is always part of good scientific conduct. In the past, technology was not sufficient to provide individual feedback, and information on the use of collected data was often provided in a general manner via websites or mass-mailing of newsletters. As a result of technological advances, it is now possible to build portals or apps to provide much more personalized information, showing a person for which purpose his/her data were used, and allowing participants to indicate whether they want to participate in specific projects or withdraw from the ongoing study. Such personalized platforms may require additional information from participants such as email address or phone number for verification purposes and provide new log-in information. The extent to which active informed consent requiring a handwritten signature is needed or it is sufficient to inform the participant and have an opt-out procedure needs to be discussed with an ethical committee any time new data collection takes place. Thinking of the different kinds of projects that will take place and the way information will be shared with participants and getting the tools ready before the start of the twin register will not only save valuable resources later on, but it may also

show the participants you will protect their rights, leading to increased trust in the twin register.

Determination of zygosity in twin registries

For a twin register, a critical measurement point is the zygosity status of a twin pair, that is, MZ or DZ, as it is the basis for subsequent research that focus on heritability estimation and genetic covariance structure modeling. It is also one of the most frequently asked questions by the twins, as they are sometimes uncertain or misinformed about their zygosity status. Even when no genetic analyses are carried out and the large datasets are used for epidemiological studies, researchers may want to correct for clustering in the data, depending on zygosity status. Misclassification of zygosity status in MZ or DZ pairs generally results in the heritability estimate going down (**Figure 1**). In extreme cases, it may even result in wrong conclusions to be drawn from variance components modeling.

Zygosity can be determined according to simple rules (see **Box 4**), but DNA testing will give the most conclusive zygosity assessment. A recent development is to genotype both twins with Single Nucleotide Polymorphism (SNP) arrays such as the Illumina Infinium Global Screening Array (GSA) or the Affymetrix Axiom World Array[25]. These arrays allow for fast genotyping for over 600,000 SNPs, which is more than is required to determine twin zygosity. However, given the reductions in genotyping costs, and the possibilities for future genetic association studies, makes a genome wide array a good investment. Of course, both twins need to provide their DNA. This can be collected by available prefabricated DNA kits for collection of buccal or saliva DNA at home, or blood can be provided at the study site. Once in the lab, DNA needs to be extracted, purified if needed, and diluted to the right concentration. The subsequent steps might be more array specific, but involve the fragmentation of the DNA into smaller pieces, then precipitation, and then hybridization to the chosen array. Here the sample fragments of DNA will 'connect' to the SNP alleles, variants of DNA sequence in humans, which are present on the array. This hybridization results in a fluorescent tag, which subsequently can be read from the array for all SNPs.

For zygosity assessment, a minimum number of typed SNPs needed is around 50; however, using between 20,000 and 30,000 typed SNPs is optimal. At the DNA level, MZ twins will share (close to) 100% of their alleles. DZ twins will share on average 50% of their alleles, similar to siblings. After using the factory standard tools for array genotyping (Beadstudio or APT-genotyper), a tool like Plink[26] can be employed to quality control the

Box 4: Basic Rules for Zygosity Determination

- Opposite-sex: DZ
- Different blood groups: DZ
- Large differences in eye, skin, and hair color: DZ
- One placenta: MZ (note that two placentas does not imply DZ)
- Alike as two peas in a pod; parents cannot tell the children apart: likely MZ
- Offspring and grandchildren cannot tell parents or grandparents and their twin apart: likely MZ
- Discordance for blood group or DNA markers: DZ

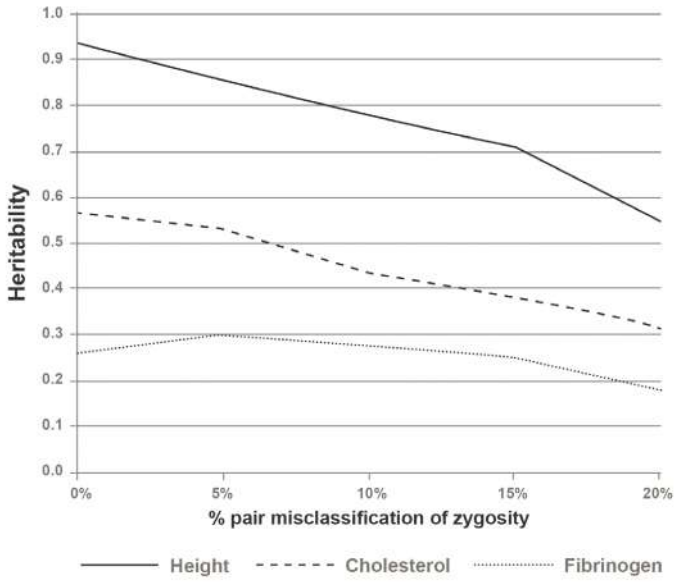


Figure 1. The effect of zygosity misclassification on the heritability estimates within a twin study

In this figure, heritability estimates for height, total cholesterol, and fibrinogen are given on the y-axis. These estimates were calculated from the phenotypic correlations 'c' between the two individuals of 391 Dutch DNA-confirmed MZ and 391 DZ pairs, with the formula (cMZ-cDZ)/(1-cDZ). Subsequently, in 5, 10, 15, and 20% of these pairs, the zygosity status was flipped from MZ to DZ, and from DZ to MZ, introducing misclassification (x-axis). Then, the heritability was re-calculated and plotted in the figure. Depending on how strong the heritability of the phenotype is, the misclassification in general reduces the overall heritability

DNA data, select an optimal number of SNPs, and determine the allele sharing in all pairs (genome option). This sharing is then given by the percentage of markers for which a pair shares no alleles (Z0), one allele (Z1), and two alleles (Z2). From these proportions, the overall sharing is calculated, by π (pi), which equals 0.5 x Z1 + Z2. Then, MZ pairs can be identified from the results by finding pairs with a $\pi > 0.90$. The DZ pairs can likewise be selected, by finding pairs that have a π between ~0.30 and ~0.70, and a similar value for sharing 1 allele (Z1) (**Figure 2**). For other values of π, researchers need to recheck which DNA sample was typed for the twins.

This approach has several more advantages. There is a useful genotyped dataset which allows for checking additional issues like genetic relatedness among participants, gender, heterozygosity, and if the study population is ethnically heterogeneous. As next steps, SNP sets can be imputed to, for example, the 1,000 Genomes phase 3 or the Haplotype Reference Consortium (HRC) genome reference panels[27]. These data can then be analyzed in genetic association studies and contribute through meta-analysis in consortia to localization of genes for complex traits, to polygenic risk scores analyses and estimation of SNP heritability, or employing Mendelian randomization to find causative relations.

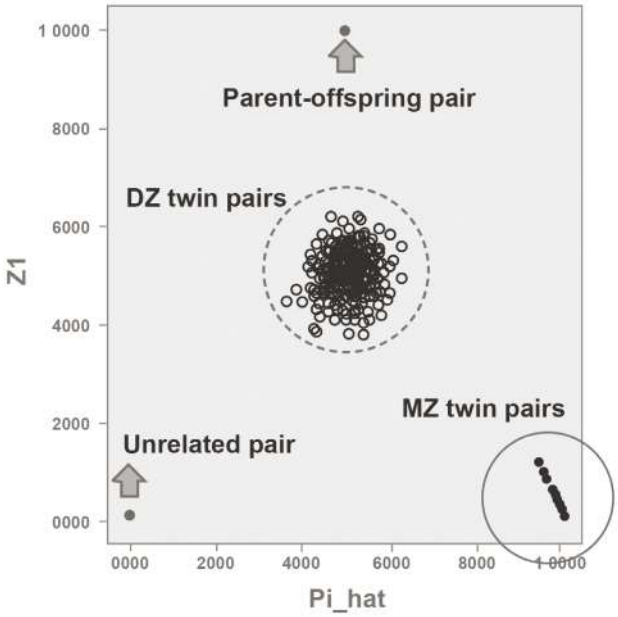


Figure 2. Allele sharing of various family pairs plotting the sharing of one allele versus Pi-hat to identify monozygotic and dizygotic twin pairs

Phenotyping: From Survey to Record Linkage

Twin registers have obtained a wide variety of phenotype data through various methods. The basic measurement method often is the survey, with registers sending out surveys at regular intervals. When deciding on what to include in a survey, the purpose of the current survey as well as the long-term goals need to be taken into account. For instance, a funded study may focus on alcohol use, but a long-term goal may be to determine how genes and lifestyle contribute to depression, so it would make sense to include a depression scale in the survey. Also, it is important to consider which data can still change over time and which data are fixed and do not need to be obtained again. This may of course be age specific. For instance, in a middle-aged group questions regarding educational level may not need to be repeated. Questionnaires also need to keep a balance between the quantity of information gathered and the participants' needs, since they should not be burdened with too many questions, risking attrition, or incorrect/missing data. While devising the first survey may seem daunting, many twin registers will be happy to share information to help the new register use well-established procedures and avoid pitfalls in survey set-up. While survey data can be obtained in all or at least large groups of participants, some data can only be collected in limited numbers. Laboratory procedures or specific phenotypes needing of complex settings, long assessment times, or expensive equipment are not easily applicable to large samples. Examples would be studies on brain imaging or extensive cognitive testing. In these cases, participants maybe invited based on specific inclusion criteria. New developments taking advantage of information technologies are modifying data collection procedures in epidemiological research and are also applied in twin studies. This includes computer-assisted surveys,

and ambulatory assessment of objective (e.g., actimetry) and self-reported (e.g., mood and exercise) phenotypes through web or mobile applications.

Data collection is not, of course, limited to surveys or laboratory assessment. The assessment of environmental exposures linked to address or workplace information and the development of exposome-wide association studies represents novel approaches to gathering information for research purposes that do not need the direct involvement of the register participant.

Record linkage to external databases (e.g., hospital, primary care, or insurance and education records) also is an invaluable source of information that has been used by registers, in Scandinavian and other countries. For example, Van Beijsterveldt et al.[28] linked phenotype information from the Netherlands Twin Register to the database of the Dutch pathological anatomy national automated archive. Record linkage was successful for over 9,000 twin pairs. The effect of chorion type was tested by comparing the within-pair similarity between monochorionic and dichorionic MZ twins on 66 traits. They concluded that the influence of the intra-uterine prenatal environment, as measured by sharing a chorion type, on MZ twin resemblance was small and limited to a few phenotypes, implying that the assumption of equal prenatal environment of mono and dichorionic MZ twins, which characterizes the classical twin design, is largely tenable.

Possibilities for Biobanking in Twin Registers

Many twin registries collect biological samples from their participants. Initially, the reason for collection of blood samples often was to have a reliable measure of zygosity based on blood group or DNA typing, but biological sample collection can also extend the potential of genetic epidemiologic research into, for example, cardiovascular and late-life health and mortality, by allowing measurement of biomarkers. Combined with the twin design, this allows estimation of the contribution of genomic factors (genetic, epigenetic, and gene expression) and biochemical factors (metabolites and proteins) to intermediate phenotypes and risk factors of disease, such as lipid levels[29]. Designs involving MZ twin pairs allow discovery of variability genes, as demonstrated for lipid levels[30]. The development of laboratory technologies has dramatically increased opportunities to study collections of bio-specimens and their related data. This allows for comprehensive studies of complex diseases and phenotypes, facilitates identification of predisposing genes and epigenetic factors, and provides support for a better understanding of disease etiology.

The organization of biobanks becomes an important element with the increase of bio-specimens and the necessity to conserve them. For example, whereas germ-line variations in the DNA sequence of a person rarely depend on the age at which a sample was collected, this is different for somatic DNA variation, epigenetic, and telomeric variation, for which the subjects' age when the specimen was collected is an important determinant[31–33]. Other determinants of epigenetic profiles are tissue/cell type[34] and lifestyle factors such as smoking. Many types of samples (e.g., whole blood) contain a mixture of cell types with distinct epigenetic profiles. In epigenetic studies of such heterogeneous samples, assessment of cell counts allows to control for variation in cellular proportions between samples.

There are multiple strategies for collection, processing, and storage of biological samples. A wide variety of specimen types may be collected and in many molecular genetic studies more than one tissue is stored, including blood and blood fractions (plasma, serum, buffy coat, and red blood cells), RNA, saliva, buccal cells, urine, hair, fecal samples, or nails. Each of these specimen types needs to be collected, processed, and stored under conditions that preserve their stability with respect to the intended future analysis[35–38].

Collection of blood specimens should be carried out by trained personnel. An evacuated tube system (vacutainers) or plastic tubes are commonly used to collect blood. Umbilical cord blood is a useful source for research purposes since the method of collection is not invasive. It can be obtained either through venous puncture of the umbilical cord or direct drainage to a sterile container immediately after delivery (vaginal or caesarean). Blood is often fractionated in components (mononuclear leukocytes, neutrophils, erythrocytes, and plasma) before being analyzed or stored. When biobanking, blood should be aliquoted across series of tubes, as most assays use only a small amount of plasma or serum and this avoids thaw/refreeze. Serum or plasma allow for analyses of classical biomarker assays, antibodies, nutrients, lipids and lipoproteins, leptin, adiponectin, growth hormone axes, thyroid axis, inflammation, liver and kidney function, innate immunity, and metabolomic and proteomic analyses.

Metabolomics is the rapidly evolving field of the comprehensive measurement of ideally all endogenous metabolites in a biological fluid. The use of mass spectrometry and nuclear magnetic resonance spectroscopy provide novel biomarkers of metabolic health[39]. Depending on the biomarker of interest, it may be important to collect, and note, whether samples were taken after fasting, how long after a meal or on a particular day, and of the menstrual cycle in women.

Whole blood, saliva or buccal cells are excellent source of DNA. Self-collection of buccal cells is a safe, simple, and cheap method that can be used to reduce the cost of specimen collection and is often preferred over blood collection by participants. Several methods are used for collecting buccal cells, including swabs, cytobrushes, and a mouthwash protocol[38,40,41]. Other sources of DNA include, for example, toe nails[42].

In contrast to DNA, RNA is very sensitive to degradation at room temperature. Transcriptomics studies require careful RNA collection, using the PAXgene Blood RNA System, which consists of a blood collection tube in which intracellular RNA is stabilized (PAXgene Blood RNA Tube) and can be isolated by using a nucleic acid purification kit (PAXgene Blood RNA Kit). Alternatively, the samples can also be by snap frozen in liquid nitrogen, or RNA can be isolated from PBMCs using Histopaque density gradients. Total RNA, including miRNA, can be isolated simultaneously from different biological sources. Plasma (300 µL) and serum isolations can be performed using miRNeasy Serum/Plasma kit from Qiagen. For isolation, after homogenization, from tissue biopsies (e.g., cartilage or adipose tissue) the miRNeasy Mini Kit from Qiagen can be used.

Many analytes, such as steroid hormones, pesticides, and a wide variety of drugs and their metabolites, can be measured in urine, making it a convenient specimen for a variety of studies. Urine collection can be performed under different conditions depending on the study goal: immediately upon rising in the morning, random urine specimens

(for drug monitoring and cytology studies), fractional specimens after the last evening meal (to compare urine analyte levels with their concentration in blood), and timed urine collections (e.g., 12 and 24 hours to allow comparison of excretion patterns). Urine specimens should be maintained on ice or refrigerated for the duration of the collection. Collection vessels are generally larger than for other liquid specimens (from 50 to 3000 mL). Due to the non-invasive method of collection and metabolic composition urine is widely used in research of metabolite biomarkers and a wide range of diseases[43].

For microbiome investigation, fecal samples can be collected easily in a sealed container following simple instructions, and their processing can provide important information for classical twin analysis, such as in the studies estimating the heritability of gut microbiota[44], and related epidemiological and molecular approaches.

Databases for Twin Registers

Both administrative processes and scientific applications require database systems that recognize the clustered structure of data collected in twin families. Administrative processes may consist of importing new participants, who may or may not be related to existing participants, address management, documenting the participation status of individuals (moved, not willing to participate, ill, and deceased), and storing information on contacts and mailings with, for example, invitations to take part in particular studies, the responses to mailings and invitations, and outcomes of approaching non-responders. Any system that keeps track of personal information needs to adhere to guidelines concerning privacy. Identifying information, such as name, date of birth and address, should be stored separately from other information collected on participants. Often, administrative and scientific processes will be supported by different database systems whose requirements depend on the dimensionality of the data. Phenotype data from surveys will require different systems than imputed genotype data which may contains as many as 50 million markers per person. Different databases may each work with separate anonymous IDs, and keys to link databases should be carefully kept.

Databases that contain multiple relatives, should consider how to store information on family relations[45,46], especially when recruitment of participants not only involves twins, but also other relatives and multigeneration pedigrees – for example, parents or offspring of twins (see **Box 5**).

Data Analyses Issues in Twin Studies: Batch Effects and Family Clustering

Phenotyping in twins has often included biomarker assessments, such as lipids or hormone levels, and increasingly include assessments obtained by means of high-throughput technologies, such as genetic variants, gene expression data, and epigenetic modifications. These data are important to understand the nature of genetic variance components as established in twin and family studies[47], and are themselves subject to such studies, for example, studies of the heritability of methylation and gene expression[15,48–52].

Subtle differences in the processing of batches of biological samples are known to give rise to batch effect. The registration of information relevant to batch (batch number, analyst, time, date) provides the means to correct for such effects, and various methods have

Box 5: Twin designs
The classical twin design encompasses MZ and DZ twin pairs but there are other designs. For instance, twin and adoption designs can be combined when twins reared apart are accessible. More often, twin registers may have the opportunity to incorporate other kind of relatives (extended twin family designs) that can be contemplated by a register even from the very beginning. These extended designs and possible combinations offer additional opportunities and statistical power to challenge research questions, such as the possibility to disentangling genetic from shared-environmental influences within family relationships[14,21,84]. The classical design may be enlarged around the twins by incorporating twins' parents (nuclear twin family design), twins' offspring (children-of-twins design), or parents, offspring, siblings, and spouses[85,86], according to available information, to finally incorporate all different kind of relationships that can be found within a register dataset. An example of such broadening of sample scope is provided by the Netherlands Twin Registry[46], which used an extended-twin pedigree, making use of all the relationship types available in their database (except teacher-student), to be able to estimate the contribution of shared household effects to neuroticism in the presence of non-additive genetic factors.

been developed to this end[53–56]. Regardless of the methods to correct for batch effects, there is agreement that it is beneficial to randomize samples evenly over batches, and that this randomization should extend to case-control status and familial relatedness[57,58]. Furthermore, sample size per batch is an important factor: the larger the sample size per batch, the more accurate the batch correction.

Batch assignment of samples collected in family members raises the question of whether samples of family members should be processed together in the same batch or should be distributed – as far as possible – over distinct batches. We examined this question in two small simulation studies (for details, see **Supplement**). In the ideal situation of a balanced allocation design with relatively large batch sample sizes, accurate correction of batch effects is feasible, as we established in a simulation study (see **Supplement**). In the first simulation study, MZ twins were selected for concordance and discordance on phenotype X, which predicted phenotype Y, where Y (e.g., a biomarker) was subject to batch effects. Given the ideal scenario of random assignment and large batch sizes, we found that allocation regime (randomized as pairs or as individuals) had little effect on the results of either the regression of Y on X, or on the twin covariance matrix of Y conditional on X. The type of correction (random effects or fixed effect correction for batch) had no bearing on these results.

In the second simulation study, we considered the decomposition of phenotypic variance into additive genetic and shared and unshared variance components (ACE model) using linear mixed modeling[59]. The sample sizes (N_{MZ} and N_{DZ}) were relatively small: $N_{MZ} = N_{DZ} = 200$ (400 pairs) or $N_{MZ} = N_{DZ}=120$ (240 pairs); the number of batches was 15 or 25. The batch assignment was random by pair (both twins share a batch) or by individual. Note that randomization by individual does not rule out batch sharing. Conditional on batch, the ACE components were 4 (A), 2 (C), and 4(E), and batch variance equaled 1 (i.e., 1/11 = 9.1% of the phenotypic variance). We conducted both one step analyses and two step analyses (correct for batch effects in step 1 and estimate variance components in step 2), and we treated the batch effects as fixed and random. The results suggested that in the one step analyses, the estimates of the variance components were as good as those

obtained in the standard ACE model (without batch effects). In the two-step analyses, we found that random assignment by individual resulted in slightly better estimates. Notably, the C variance components was underestimated following random assignment by pairs (see **Supplement**).

Note that in the absence of batch effects, family clustering may still be an issue in statistical inference, based on the assumption that the data are independently and identically distributed. For instance, in genome-wide association studies (GWAS), family clustering violates the independence assumption. Happily, family clustering does not pose any statistical problems, as random effects modeling and generalized estimating equations can be used to either accommodate or to correct for the effects of family clustering, or more generally for genetic relatedness[60–62]. Regardless of randomization scheme (or not), detailed information should be recorded on batch (date and time of processing), operator (technician), plate number, and position (row and column).

Retaining the Twins

To retain participants in a longitudinal study, it is important to remain in contact. Many twin registers have set up a website providing information of the latest study results, news on grants obtained, PhDs awarded, and more general information on twin meetings and such. However, these may not be the best ways to form an actual connection between the twin register and participants. Most twin registers therefore also contact their participants in a more personal manner, either by letter or e-mail, sending out a regular newsletter to make the participant aware that the register is still seeing them as a valuable contributor. A number of twin registers also organize events in which twins and their family members can meet each other but also can meet the researchers and ask any questions they may have in person. Worth mentioning here is the annual gathering of twins at the Twins Day Festival in Twinsburg Ohio, where researchers are welcome to recruit twins for specific studies. Unfortunately, financial limitations generally prevent the twin registers from organizing such large and regular gatherings, but when meetings are organized, they are generally judged as very valuable.

In addition to general information, personalized information may also be given out to participants. When participants take part in specific projects, information on test scores (e.g., the results of an IQ tests or the cholesterol levels obtained in a blood sample) may be returned to the participants, accompanied by an explanation of the results. However, often little feedback is provided to participants related to the surveys completed during the longitudinal follow-up, due to the material and personnel costs needed for sending personalized reports to the large number of participants generally included in a twin register. However, as technology advances new ways emerge of providing personalized information. Participants' portals may provide individual reports without needing to write and post separate reports. At the Netherlands Twin Register, such an effort is now well underway, with participants obtaining information on the survey results via the MyNTR portal. As with the informed consent, it is important to consider the requirements of such a portal in advance. Constructing a participant panel even before starting the actual twin register that includes a number of twins who are willing to think about the various aspects involved in providing feedback would be helpful in setting this up in the best way possible for twins and the register support staff.

Conclusion

Twin registers have a long and successful history and a brilliant future as a research resource. The uniqueness of twin samples, the soundness and diversity of the methodological approaches, and the huge amount of data accumulated during the last decades characterize twin registers as invaluable contributors to the advancement of science, including social science. Their versatility to adapt to multiple scenarios and their orientation to collaborative work will preserve their value in the future as priceless instruments for the expansion of knowledge in the complexities of human phenotypes.

Although the global research agenda in the coming decades is difficult to forecast, twin registers can contribute to our understanding in virtually all areas related to human health and behavior. Population-based registers, especially when representative of the general population, are still cohorts of enormous epidemiological interest. The unique characteristics of twin studies, including the ability to control both genetic and shared environmental background, allow for addressing questions that are not easily solved in any other research design. These capacities make them extremely useful for gene-environment transaction research or causal inference studies[63,64]. Twin pairs — in particular those that are MZ — are remarkably informative in respect to variability of phenotypic expression, pathogenic mechanisms, epigenetics, and post-zygotic mutagenesis, and may serve as a model for research on genetic defects[15,18,65,66]. Participation of twins in co-twin, control-designed, and randomized controlled trials is an informative, albeit infrequently used, design[67]. The use of twin studies has been advocated for guiding post-GWAS studies on the effects associated with genetic variants[68], enabling stronger tests of causal hypotheses[69], formulating future strategies in pharmacogenomics research[70], or refining phenotypic definitions and evaluating biomarkers for disease[15]. Furthermore, due to their amenability to numerous nonclassical study designs, data based on twin registers can integrate with other resources to boost research in virtually every field of human research. Probably the best example is provided by the participation of twin biobanks in many of the large association studies (GWAS and EWAS) that have been published in the last decade.

An additional feature empowering twin registers relies on their orientation to collaborative work. The community of twin registers has a long history of successful alliances. The very nature of their origin as research resources and their scientific environment imply, on the one hand, the existence of matching data across different registers and, on the other hand, the need for very large samples in order to find answers to some of the research questions investigators are interested in. In these circumstances, collaboration is not only practicable, but it is a must. Multiple consortia and collaboration initiatives have seen the light as an answer to those needs. The GenomeEuTwin[71], EuroDiscoTwin[72], or the CODATwins (COllaborative project of Development of Anthropometrical measures in Twins)[73] consortia are just a few examples of associative efforts, joining together data from a large number of twin cohorts in order to advance in the analysis of the genetic and environmental underpinnings of human complex phenotypes. Other initiatives, such as the International Network of Twin Registries (INTR)[74] have emerged from the International Society for Twin Studies, aiming to foster collaboration and serve as a platform for networking and establishing research relationships between twin registers and between them and the global scientific community.

These collaborative efforts have a parallel outcome on infrastructures related to the registers, such as biobanks. In the same way that registers multiply their scientific impact when joining efforts, the effective use of biobank resources depends on their accessibility. Building a centralized database for the research community allows storing of raw and processed data, reference data for case-control studies and imputation, and linking to clinical phenotypes, so that data can be effectively used not only by single research groups, but also in collaborative multi-center and consortium projects. For instance, the advent of the GWAS method took advantage of such multi-center collaborations in order to lead to the successful identification of thousands of variants that are robustly associated with complex disease phenotypes. The big databases permit research on genetic, methylation, expression level, available protein, lipid, metabolite level information, and on disease/phenotype level. In Europe, for instance, a range of biobanks joined in the Biobanking and BioMolecular Resources Research Infrastructure and national hubs (e.g., www.bbmri.nl) generated omics data by the same platforms and shared these combined with existing phenotype data.

Nowadays, the advancement of scientific knowledge requires such collaborations to gain explanatory power and optimize the invested resources. Twin registers, and associated biobanks, have an enormous potential that multiply when joining efforts, and new or growing registers are always welcome to this endeavor. In this article, we have outlined what we feel are the main principles and recommendations for the establishment and management of a twin register, from its inception to its actual development. As pointed out before, our intention has not been to enumerate a detailed checklist of actions, or a complete step-by-step technical guide on this process, but rather to highlight the main aspects that, from our perspective, need to be taken into account for being able to make the difference between an isolated initiative and a successful long-lasting scientific resource.

References

1. Galton, F. The history of twins, as a criterion of the relative powers of nature and nurture. *Fraser's Magazine* 566–576 (1875).
2. Fisher, R. A. The Genesis of Twins. *Genetics* **4**, 489–499 (1919).
3. Thorndike, E. L. *Measurement of Twins.* (The Science Press, 1905).
4. Rende, R. D., Plomin, R. & Vandenberg, S. G. Who discovered the twin method? *Behavior genetics* **20**, 277–285 (1990).
5. Poll, H. Über Zwillingsforschung als Hilfsmittel menschlicher Erbkunde LK. *Zeitschrift für Ethnologie* – **46**, 87–105 (1914).
6. Merriman, C. *The intellectual resemblance of twins.* (Psychological Review Co., 1924).
7. Siemens, H. W. *Die Zwillingspathologie (Twin pathology: its meaning, its method, results so far).* (Springer Verlag, 1924).
8. Holzinger, K. J. The relative effect of nature and nurture influences on twin differences. *Journal of Educational Psychology* **20**, 241–248 (1929).
9. Mayo, O. Early research on human genetics using the twin method: who really invented the method? *Twin research and human genetics : the official journal of the International Society for Twin Studies* **12**, 237–245 (2009).
10. Lejeune, J., Turoin, R. & Gautier, M. [Mongolism; a chromosomal disease (trisomy)]. *Bulletin de l'Academie nationale de medecine* **143**, 256–265 (1959).
11. Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. *Nature reviews. Genetics* **3**, 872–882 (2002).
12. Martin, N., Boomsma, D. & Machin, G. A twin-pronged attack on complex traits. *Nature genetics* **17**, 387–392 (1997).
13. Martin, N. G., Eaves, L. J., Kearsey, M. J. & Davies, P. The power of the classical twin study. *Heredity* **40**, 97–116 (1978).
14. Posthuma, D. & Boomsma, D. I. A note on the statistical power in extended twin designs. *Behavior genetics* **30**, 147–158 (2000).
15. Van Dongen, J., Slagboom, P. E., Draisma, H. H. M., Martin, N. G. & Boomsma, D. I. The continuing value of twin studies in the omics era. *Nature Reviews Genetics* **13**, 640–653 (2012).
16. Ayorech, Z. *et al.* Publication Trends Over 55 Years of Behavioral Genetic Research. *Behavior genetics* **46**, 603–607 (2016).
17. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature genetics* **47**, 702–709 (2015).
18. Bell, J. T. & Spector, T. D. A twin approach to unraveling epigenetics. *Trends in genetics: TIG* **27**, 116–125 (2011).
19. Groen-Blokhuis, M. M., Middeldorp, C. M., van Beijsterveldt, C. E. M. & Boomsma, D. I. Evidence for a causal association of low birth weight and attention problems. *Journal of the American Academy of Child and Adolescent Psychiatry* **50**, 1247–54.e2 (2011).
20. Kaprio, J. & Silventoinen, K. Advanced methods in twin studies. *Methods in molecular biology (Clifton, N.J.)* **713**, 143–152 (2011).
21. Knopik, V. S., Neiderhiser, J., DeFries, J. C. & Plomin, R. *Behavioral genetics.* (Worth Publishers, 2017).

22. Goldberg, J. *et al.* Identification of a cohort of male and female twins aged 65 years or more in the United States. *American journal of epidemiology* **145**, 175–183 (1997).

23. Bechger, T. M., Boomsma, D. I. & Koning, H. A Limited Dependent Variable Model for Heritability Estimation with Non-Random Ascertained Samples. *Behavior Genetics TA – TT –* **32**, 145–151 (2002).

24. Neale, M. C. & Eaves, L. J. Estimating and controlling for the effects of volunteer bias with pairs of relatives. *Behavior genetics* **23**, 271–277 (1993).

25. Ehli, E. A. *et al.* A method to customize population-specific arrays for genome-wide association testing. *European journal of human genetics: EJHG* **25**, 267–270 (2017).

26. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* **81**, 559–575 (2007).

27. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics* **48**, 1279–1283 (2016).

28. Van Beijsterveldt, C. E. M. *et al.* Chorionicity and Heritability Estimates from Twin Studies: The Prenatal Environment of Twins and Their Resemblance Across a Large Number of Traits. *Behavior genetics* **46**, 304–314 (2016).

29. Snieder, H., van Doornen, L. J. & Boomsma, D. I. The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *American journal of human genetics* **60**, 638–650 (1997).

30. Berg, K. Variability gene effect on cholesterol at the Kidd blood group locus. *Clinical Genetics* **33**, 102–107 (1988).

31. Fraga, M. F. *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10604–10609 (2005).

32. Slagboom, P. E., Droog, S. & Boomsma, D. I. Genetic determination of telomere size in humans: a twin study of three age groups. *American journal of human genetics* **55**, 876–882 (1994).

33. Van Dongen, J. *et al.* Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nature Communications* **7**, 11115 (2016).

34. Finnicum, C. T. *et al.* Relative Telomere Repeat Mass in Buccal and Leukocyte-Derived DNA. *PloS one* **12**, e0170765 (2017).

35. García-Closas, M. *et al.* Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **10**, 687–696 (2001).

36. Holland, N. T., Smith, M. T., Eskenazi, B. & Bastaki, M. Biological sample collection and processing for molecular epidemiological studies. *Mutation research* **543**, 217–234 (2003).

37. Tworoger, S. S. & Hankinson, S. E. Collection, processing, and storage of biological samples in epidemiologic studies: sex hormones, carotenoids, inflammatory markers, and proteomics as examples. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **15**, 1578–1581 (2006).

38. Vaught, J. B. & Henderson, M. K. Biological sample collection, processing, storage and information management. *IARC scientific publications* 23–42 (2011).

39. Suhre, K. *et al.* Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. *PloS one* **5**, e13953 (2010).

40. Meulenbelt, I., Droog, S., Trommelen, G. J., Boomsma, D. I. & Slagboom, P. E. High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *American journal of human genetics* vol. 57 1252–1254 (1995).

41. Min, J. L. *et al.* High microsatellite and SNP genotyping success rates established in a large number of genomic DNA samples extracted from mouth swabs and genotypes. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **9**, 501–506 (2006).

42. Hogervorst, J. G. F. *et al.* DNA from nails for genetic analyses in large-scale epidemiologic studies. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* vol. 23 2703–2712 (2014).

43. Duarte, I. F., Diaz, S. O. & Gil, A. M. NMR metabolomics of human blood and urine in disease research. *Journal of pharmaceutical and biomedical analysis* **93**, 17–26 (2014).

44. Goodrich, J. K. *et al.* Genetic Determinants of the Gut Microbiome in UK Twins. *Cell host & microbe* **19**, 731–743 (2016).

45. Boomsma, D. I. *et al.* Design and implementation of a twin-family database for behavior genetics and genomics studies. *Twin research and human genetics : the official journal of the International Society for Twin Studies* **11**, 342–348 (2008).

46. Boomsma, D. I. *et al.* An Extended Twin-Pedigree Study of Neuroticism in the Netherlands Twin Register. *Behavior genetics* **48**, 1–11 (2018).

47. Schadt, E. E. *et al.* An integrative genomics approach to infer causal associations between gene expression and disease. *Nature genetics* **37**, 710–717 (2005).

48. Bell, J. T. & Spector, T. D. DNA methylation studies using twins: what are they telling us? *Genome Biology* **13**, 172 (2012).

49. McRae, A. F. *et al.* Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome biology* **15**, R73 (2014).

50. Petronis, A. Epigenetics and twins: three variations on the theme. *Trends in Genetics* **22**, 347–350 (2006).

51. Petronis, A. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* **465**, 721–727 (2010).

52. Wright, F. A. *et al.* Heritability and genomics of gene expression in peripheral blood. *Nature genetics* **46**, 430–437 (2014).

53. Chen, C. *et al.* Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. *PloS one* **6**, e17238 (2011).

54. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics (Oxford, England)* **8**, 118–127 (2007).

55. Lazar, C. *et al.* Batch effect removal methods for microarray gene expression data integration: a survey. *Briefings in bioinformatics* **14**, 469–490 (2013).

56. Sun, Z. *et al.* Batch effect correction for genome-wide methylation data with Illumina Infinium platform. *BMC medical genomics* **4**, 84 (2011).

57. Leek, J. T. *et al.* Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature reviews. Genetics* vol. 11 733–739 (2010).

58. Nygaard, V., Rødland, E. A. & Hovig, E. Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. *Biostatistics (Oxford, England)* **17**, 29–39 (2016).

59. McArdle, J. J. & Prescott, C. A. Mixed-effects variance components models for biometric family analyses. *Behavior genetics* **35**, 631–652 (2005).

60. Li, X., Basu, S., Miller, M. B., Iacono, W. G. & McGue, M. A rapid generalized least squares model for a genome-wide quantitative trait association analysis in families. *Human heredity* **71**, 67–82 (2011).

61. Lippert, C. *et al.* FaST linear mixed models for genome-wide association studies. *Nature methods* **8**, 833–835 (2011).

62. Minică, C. C., Dolan, C. v, Kampert, M. M. D., Boomsma, D. I. & Vink, J. M. Sandwich corrected standard errors in family-based genome-wide association studies. *European journal of human genetics: EJHG* **23**, 388–394 (2015).

63. Johnson, W., Turkheimer, E., Gottesman, I. I. & Bouchard, T. J. J. Beyond Heritability: Twin Studies in Behavioral Research. *Current directions in psychological science* **18**, 217–220 (2010).

64. McGue, M. The end of behavioral genetics? 2008. *Behavior genetics* **40**, 284–296 (2010).

65. Castillo-Fernandez, J. E., Spector, T. D. & Bell, J. T. Epigenetics of discordant monozygotic twins: implications for disease. *Genome Medicine* **6**, 60 (2014).

66. Zwijnenburg, P. J. G., Meijers-Heijboer, H. & Boomsma, D. I. Identical but not the same: The value of discordant monozygotic twins in genetic research. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **153B**, n/a-n/a (2010).

67. Sumathipala, A. *et al.* Twins as Participants in Randomized Controlled Trials: A Review of Published Literature. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **21**, 51–56 (2018).

68. Dick, D. M. *et al.* Post-GWAS in Psychiatric Genetics: A Developmental Perspective on the "Other" Next Steps. *Genes, brain, and behavior* **17**, e12447 (2018).

69. Iacono, W. G. *et al.* The utility of twins in developmental cognitive neuroscience research: How twins strengthen the ABCD research design. *Developmental cognitive neuroscience* **32**, 30–42 (2018).

70. Rahmioğlu, N. & Ahmadi, K. R. Classical twin design in modern pharmacogenomics studies. *Pharmacogenomics* **11**, 215–226 (2010).

71. Peltonen, L. GenomEUtwin: a strategy to identify genetic influences on health and disease. *Twin research: the official journal of the International Society for Twin Studies* **6**, 354–360 (2003).

72. Willemsen, G. *et al.* The Concordance and Heritability of Type 2 Diabetes in 34,166 Twin Pairs From International Twin Registers: The Discordant Twin (DISCOTWIN) Consortium. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **18**, 762–771 (2015).

73. Silventoinen, K. *et al.* The CODATwins Project: The Cohort Description of Collaborative Project of Development of Anthropometrical Measures in Twins to Study Macro-Environmental Variation in Genetic and Environmental Effects on Anthropometric Traits. *Twin research and human genetics: the official journal of the International Society for Twin Studie*s **18**, 348–360 (2015).

74. Buchwald, D. *et al.* International network of twin registries (INTR): building a platform for international collaboration. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **17**, 574–577 (2014).

75. Sukhanov S. O psihozah u bliznetsov [On psychosis in twins]. *Klinicheskij jurnal* 341–352 (1900).

76. Yudin, T. I. O shodstve psihoza u bratjev i sester. [On similarity of psychosis in brothers and sisters]. *Sovremennaya psihiatria* 337–342 (1907).

77. Levit, S. G. Twin Investigations in the U. S. S. R. *Journal of Personality* **3**, 188–193 (1935).

78. Grigorenko, E. L. & Ravich-Shcherbo, I. v. Russian psychogenetics: sketches for the portrait. in *Psychology of Russia: past, present, future* (eds. Grigorenko, E. L., Ruzgis, Patricia. & Sternberg, R. J.) 84–121 (Nova Science Publishers, 1997).

79. Malykh, S., Egorova, M. & Meshkova, T. *Osnovi povedencheskoj genetiki [Foundation of Behavioral Genetics].* (Epidaurus, 1998).

80. Lykken, D. T. Presidential address, 1981. *Research with twins: the concept of emergenesis. Psychophysiology* **19**, 361–373 (1982).

81. Martin, N. G. & Martin, P. G. The inheritance of scholastric abilities in a sample of twins. I. Ascertainments of the sample and diagnosis of zygosity. *Annals of human genetics* **39**, 213–218 (1975).

82. Martin, N. G. The inheritance of scholastic abilities in a sample of twins. II. Genetical analysis of examinations results. *Annals of human genetics* **39**, 219–229 (1975).

83. Neale, M. C., Eaves, L. J. & Kendler, K. S. The power of the classical twin study to resolve variation in threshold traits. *Behavior genetics* **24**, 239–258 (1994).

84. Keller, M. C., Medland, S. E. & Duncan, L. E. Are extended twin family designs worth the trouble? A comparison of the bias, precision, and accuracy of parameters estimated in four twin family models. *Behavior genetics* **40**, 377–393 (2010).

85. Keller, M. C. *et al.* Modeling extended twin family data I: description of the Cascade model. *Twin research and human genetics : the official journal of the International Society for Twin Studies* **12**, 8–18 (2009).

86. Maes, H. H., Neale, M. C. & Eaves, L. J. Genetic and environmental factors in relative body weight and human adiposity. *Behavior genetics* **27**, 325–351 (1997).

Supplement

Processing biological material in batches may give rise to batch effects, i.e., intra-batch correlation greater than zero. A question that is specific to the twin design (or any other design with naturally clustered observations) concerns the manner of allocation of twins to batches. One may allocate randomly by individual twin, or randomly by twin pair. The latter implies that the twin pairs share the batch, the former does not rule out batch sharing. The following are the results of simulation studies carried out to answer this question in three situations.

How to allocate twins to batch in assay of metabolites in an extremely discordant and concordant (EDAC) twin design?

Discordant and concordant twin pairs are selected on the basis of phenotypic scores, for example aggression scores, for a biomarker study. Assays on the twins' urine samples are done to measure metabolites. The aim is to determine the association between metabolite levels and aggression. The metabolites are determined on plates (i.e., in batches). The present question concerns the allocation of twins to batch, given that plate is a source of systematic variation:

1) assign twin pairs randomly to batches

2) assign twin members (individuals) randomly to batches.

An additional question, specific to the EDAC design, is the choice of the independent variable. As the selection is on aggression scores it is statistically expedient to regress metabolite (predictor) on aggression (dependent). Selection on the predictor does not affect the regression, and if the selection is based on an EDAC scheme, the selection results in little loss of power. Alternatively, one may choose to regress metabolite on the binary aggression scores (e.g., 0=low, 1=high). Regression on the continuous score is expected to confer greater power.

We make the following assumptions concerning the analysis. We assume that the twin data are to be analyzed in a single statistical model, which will include the discordant and concordant twins. With respect to this model, in testing the association of metabolite and aggression, we have to accommodate 1) the inherent two-level structure (family clustering of twins in twin pairs), and 2) the batch effects. We consider two models:

1) Linear mixed model, in which effect of batch is accommodated by means of a random effect (variance component).

2) Fixed regression model with metabolite corrected for batch in one or two step procedure. Two step procedure: regress metabolite on plate first, use residuals in regression on predictor. One step procedure: regress metabolite on plate and on predictor at the same time.

The association between metabolite and aggression is accommodated by means of a fixed effect, i.e., regression of metabolite on binary (0/1) or continuous aggression score.

Simulation 1. Random effects model.

The metabolite explains 5% of the variance in aggression. The heritability of metabolite is .6, the heritability of the residual of aggression is 5. The number of batches is 70, the number of twin pairs is 600. The true phenotypic variances of metabolite and aggression are both set to equal 1 and the variance is .25. All variables have zero mean. So, the metabolite variance is 1.25. The number of replications is 50.

Allocation regime (pairs vs. individuals) has no effect on the estimate of the parameter of interest. We note that, as expected, regression on continuous predictors confers more power than regression on binary predictor (0/1). The variance components (Additive genetic, Environment and Batch) appear to be slightly downwards biased in the allocation by pair, but accurate in the allocation by individual condition.

Simulation 2: fixed batch effects in two steps or one step.

It may be expedient to carry out analyses in two steps, i.e., first correct for batch effects, and second carry out the analysis of actual interest. We compared one and two step analyses in simulation 2. We used linear mixed modeling in simulation 1 (estimating the twin covariance conditional on predictor and batch). Here we use GEE (generalized estimating equations), i.e., we correct the standard errors after the analyses using a sandwich correction.

One-step: using GEE regress metabolite on predictor and batch simultaneously.

Two-steps: first correct metabolite for batch effect and then use GEE to analyze the residuals.

Conclusions are the same as those based on Simulation 1. Allocation regime (pairs vs. individuals) has little effect on the test of the parameter of interest. Again, as expected, regression of continuous predictors confers more power that regression on binary predictor (0/1). We see little differences between one and two step procedure.

Simulation 3: More extreme selection and fixed plate effects in two step or one step.

This simulation is the same as simulation 2. However here we employ a more extreme selection criterion rather than a mean split (Simulations 1 and 2), the selection of high and low scoring twins is on the basis of the criteria >.5 std unit or < -.5 std units. As in simulation 2, we carry out one and two step analyses using GEE. Given the selection, we set the total sample size to 5000 (random sample) and select from this sample based on the criteria mentioned.

The conclusions are consistent with those of simulations 1 and 2. The allocation regime, i.e., pairs vs. individuals, has little effect on the test of the parameter of interest. The regression of continuous predictors confers more power that regression on binary predictor (0/1), as expected. There is little difference between the results of the one and two step procedure.

How to allocate twins to batches in assay of metabolites in the classical twin design

Simulation 4: estimating genetic and non-genetic variance components in the twin design.

In simulation 1, we noted that batch allocation by twin pair appeared to result in a slight bias in the estimates of the variance components. In simulation 4, we examined the effect on variance components by fitting an ACE model to twin data. Here we treat batch as a random and as a fixed effect, and we carry out both one step and two step analyses. We consider relatively small sample sizes. We use linear mixed modeling with REML (restricted maximum likelihood) estimation.

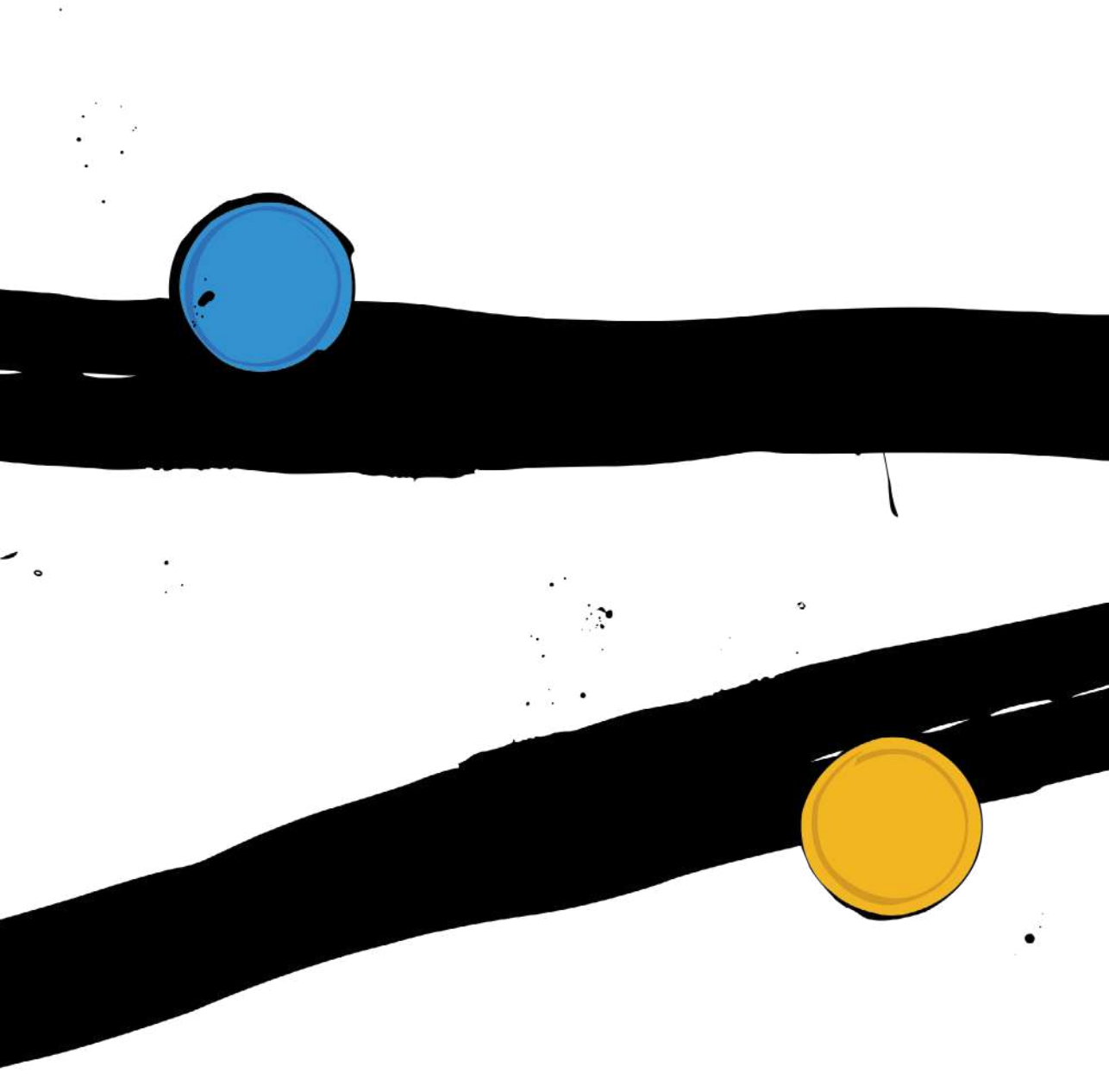
The results demonstrate that allocation regime has little effect in the one step analyses, regardless of whether this is based on random effects or fixed effects modeling of batch. In the two step procedures, we note a downward bias in the estimates of the A (additive genetic) and C (common environment) variance components. This bias is greater in given the allocation by twin pairs, and greater as the number of batches increases.

The supplementary materials are available online at

https://doi.org/10.1017/thg.2018.23

Chapter 10

Twin-Singleton comparisons across multiple domains of life



Abstract

In this chapter we address the question whether individuals born from a multiple pregnancy differ from singletons. The answer to this question is important for health care professionals and researchers, as well as multiples themselves and their family members.

First, we review findings from the literature with respect to twin – non-twin differences in early life and conclude that a multiple pregnancy increases the risk of congenital problems and mortality for the unborn and newborn children.

Next, we provide an overview of the outcomes of comparing adult twins to their singleton siblings across a wide range of traits assessed in the Netherlands Twin Register (NTR). In a within-family design, comparing twins to siblings from the same family, we correct for familial confounding. Overall, hardly any evidence was found for the presence of twin-sibling differences for the five domains explored, which included body composition and physical development, personality and psychopathology, behavioral and sociodemographic traits, physiological parameters and physical disease, and cognitive function. With the exception of minor differences in body composition, twins do not seem to differ from singletons, when taking family factors into account.

In conclusion, while being a twin can be seen as special, adult twins are similar to ordinary siblings across most domains of life.

Published as: Willemsen, G., Odintsova, V., de Geus, E. & Boomsma, DI. (2021).Twin-Singleton comparisons across multiple domains of life. In: *Twin and Higher-order Pregnancies* (eds. Khalil A., Lewi L. & Lopriore E.). Springer International Publishing, pp. 51-71.

Introduction

Over the past decades, twin and higher-order multiple pregnancy rates have increased in many countries worldwide[1–3]. While the increased number of infertility treatments is often stated as the reason for this increase, the increase in maternal age is also a significant contributor[4]. Compared to singletons, that is children born from a single pregnancy originating with a single zygote, children born from a multiple pregnancy are born on average 3 weeks earlier[5], and are about 4 cm shorter and 1 kilo lighter at birth[6]. Multiples, that is twins, triplets and higher-order multiples, may be at increased risk of complications during pregnancy, delivery and in the postnatal period. The fetal origins hypothesis and the developmental origins of health and disease hypothesis, often referred to as the Barker hypothesis, poses that low birthweight reflects intra-uterine growth retardation that may affect later development and increases the risk of disease development and earlier mortality[7]. Numerous studies in the general population have indeed found evidence for an association of low birthweight with diseases including hypertension, cardiovascular disease, and diabetes[8], and with all-cause mortality[9]. An important question is whether this hypothesis extends to twins, with twins compared to singletons having a higher risk of disease development and early mortality due to their on average lower birthweight, which would have consequences for the health care of multiples.

Twins do not only differ from singletons in birthweight. They also shared a womb, and grew up with a sibling who is the exact same age and often the same sex and even, depending on the zygosity, alike in physical appearance. Growing up with such a close companion may lead to differences in behavior as it may encourage or discourage certain behaviors and lifestyles through mechanisms of social interaction[10]. Twins also tend to have older parents than non-twin children, and for a range of behavioral and cognitive outcomes the children of older parents do somewhat better than children of younger fathers and mothers[11,12].

The above raises the important question of whether twins are similar to singletons, both with regard to health and non-health related traits. The answer to this question is relevant to health care professionals but also to researchers making use of the twin design in studies to unravel the influence of genetic and environmental factors on traits of interest. The generalization of findings based on twin research rests on the assumption that twins do not differ from non-twin individuals. For example, in studies of gene-by-environment interaction, twins should not differ from singletons regarding their exposure to the environment under study. While this seems a reasonable assumption, for example in case of the death of parents, this may be different for some other life events such as divorce or having an intimate relationship with others, because twin relations may discourage such relationships[13–15]. Importantly, in addition to healthcare professionals and scientists, twins themselves and their family members are interested in the question whether and how twins are different from singletons.

In this chapter, we first discuss findings from the literature with respect to twin – non-twin differences in early life, e.g. birthweight, pre- and perinatal mortality differences, and congenital problems. In the second part, we provide the results for adult twin – non-twin comparisons across a wide range of traits assessed in the Netherlands Twin Register (NTR), including a large series of biomarkers.

Congenital Disorders and Infant Mortality

When the proportion of multiple births in Europe increased from 1.9% (1984-88) to 3.1% (2004-2008), the prevalence of congenital anomaly from multiple births also increased[16]. A multiple pregnancy carries extra risk for fetuses and neonates. Twins grow slower during the third trimester than singletons[17], experience more intrauterine growth restriction[18] and are more likely to have a low birthweight[19]. This may influence both twins but may also be limited to one of the members in a twin pair, in the case of selective fetal growth restriction[20]. Compared to singletons, multiple birth children show a substantially higher rate of overall perinatal mortality[21,22] and stillbirth[23,24], though some studies report lower perinatal mortality rates in preterm twin pregnancies, possibly due to increased medical surveillance in the case of a multiple pregnancy[25,26].

In about 60% of twin pregnancies, malpresentation occurs, with one or both of the twins not optimally positioned for birth[27]. Surgical delivery and assisted interventions during vaginal delivery are common in multiple pregnancies[22,28,29] and multiple birth is a risk factor for low Apgar scores[30,31]. While preterm delivery and low birthweight explain part of the higher perinatal mortality and morbidity rates in twins, the risk of adverse outcomes is still higher when comparing twins with normal birthweight to singletons of the same birthweight[2].

The second-born twin generally faces the greatest risks as obstetric complications, such as placental separation, cord prolapse, uterine atony, prolonged intertwin delivery time, and cervical spasm, may occur after delivery of the first-born twin[29] and can cause fetal distress, low Apgar scores, and neonatal morbidity[32,33].

Prematurity and low birthweight are also associated with neurodevelopmental disorders and cerebral injury. Cerebral palsy is reported five to ten times more often in twins compared with singletons[34–36]. The risk of cerebral palsy is affected by birth asphyxia that causes cerebral impairment[36] and by neonatal death or stillbirth in the co-twin[37].

There is considerable evidence that babies from multiple pregnancies have a higher risk of total congenital anomalies than singleton babies[36,38–40], with reports of a relative risk of 1.29 for congenital anomaly in multiple births relative to singletons[1]. Hall[41] estimated that probably 10% of monozygotic twins are born with a congenital anomaly. The most common anomalies in twins and singletons for which twins have a higher risk than singletons are cardiovascular anomalies[38,42–46]. Higher rates in twins are also reported for anomalies of the central nervous system, the digestive system, in particular gut atresias, the genito-urinary track and musculoskeletal systems[38,43]. In addition, neural tube defects have been reported more often in twins[47] and the prevalence of clubfoot is twice that of the general population[48]. Twins, however, do not seem to have an excess risk of oral cleft compared to singletons[48,49]. With respect to chromosomal abnormalities, study results differ. Some have shown lower rates of chromosomal abnormalities in twins compared to singletons[39,42,46,50,51], while others showed no differences[38,43]. Some chromosomal anomalies and imprinting disorders are more prevalent in monozygotic twins with discordant presentation. For example, in Beckwith-Wiedemann syndrome the majority of affected twins have an unaffected monozygotic co-twin who may have only some features of the disease, and it has been suggested that a methylation failure in the twinning process is involved[52].

Several explanations have been proposed for the higher rates of congenital anomalies and malformations in twins, including disturbances in early embryonic development, especially in monozygotic twinning[38,42,53], hemodynamic instability in monozygotic placentation[54], contribution of artificial reproductive technologies and other treatments of infertility[16,38,55], as well as maternal age at pregnancy[56]. In the majority of the cases, congenital anomalies occur in discordant pairs where only one twin is affected[57]. The etiology is poorly understood but may involve epigenetic factors[58] as was found in a study of monozygotic twin girls who were discordant for a caudal duplication anomaly[59]. The coding region of the *AXIN1* was sequenced in both twins and while no mutation was detected, this region was significantly more methylated in the affected twin than in the unaffected twin.

Several conditions are unique to multiple pregnancies such as monochorionic-monoamniotic condition, twin-twin transfusion syndrome (TTTS), and some rare malformations such as conjoined twins, *fetus in fetu,* and acardiac malformation[60,61]. Many congenital anomalies in twins are more common in monozygotic twins than in dizygotic twins[39] and within the monozygotic twins, more common in monochorionic than dichorionic twins[38]. Together with TTTS, congenital anomalies are an additional risk for mortality and adverse neurodevelopmental outcome in monochorionic twins[62–64].

A very rare condition, which may occur in monozygotic twins, involves the reversal of the internal organs known as *situs inversus partialis or totalis[65].* Several case reports of monozygotic twins with situs inversus are reported in the literature[66–70]. This may be related to the phenomenon of “mirror twins”, when the features appear asymmetrical in co-twins. For example, left-handedness in one twin and right-handedness in the other twin may be the expression of an anatomical mirror image at the level of the nervous system[68]. Other explanations include conjoined twinning, a late division of the embryo leading to monozygotic twinning, and a malrotation of the viscera during early embryonic life[65,68].

In conclusion, a multiple pregnancy increases the risk of congenital problems and mortality of the unborn and newborn. Still, the majority of twins are born healthy in countries with good healthcare systems and develop normally. This however, does not exclude the possibility that twins differ from singletons in more subtle ways. In the next part, we therefore explore possible differences between twins and their singleton siblings.

Adult twin-sibling comparisons across a wide variety of traits

Background and Procedure

While a twin pregnancy carries, as described above, a number of risks, the majority of twins are born healthy. One important question is whether such twins differ from singletons in their development and health at later age. To provide insight into the potential differences between multiples and singletons, it is essential to choose the correct reference group to which to compare the multiples. By selecting a group of singletons from the general population, this may introduce a bias as this population would also include individuals from one-child families, confounding the effects of being born after a singleton pregnancy with those of having no siblings. Even when choosing singletons from families with

more than one child, this does not correct for potential differences across twin and non-twin families, such as parental behaviors or parental genotypes. Hence, differences observed at the population level between twins and non-twins in so-called between-family comparisons should be interpreted carefully. Differences may reflect true effects, but they may also be confounded by between-family differences in, e.g., family structure, urban-rural residency, and multiple other factors, including the maternal genotype, which is known to be associated with DZ twinning[71].

One way of eliminating these problems is by comparing twins to their own singleton siblings. This design optimally matches controls (siblings) and cases (twins), as twins and siblings come from the same family, and largely share their genetic background and family environment[72]. Within-family designs are becoming common in molecular genetics, where it is recognized that gene-outcome associations found among unrelated individuals may reflect between-family variation in genetic and environmental factors. A within-family comparison reduces confounding by these factors but does require statistical approaches that take into account the dependencies in the data such as paired-sample tests. Alternatively, differences due to genetic and environmental factors can be assessed in genetic structural equation models that simultaneously model the mean and the covariance structure in the data[73]. This approach is often taken in behavioral genetic studies, when the aim of the study is to estimate genetic and non-genetic contribution to the observed variance in a trait of interest. These studies, however, do not always report the outcomes of twin-sibling comparisons and sometimes assume that twins are similar to non-twin siblings.

In the following, we employed the within-family design in an adult sample from the Netherlands Twin Register. The Netherlands Twin Register (NTR) is one of the larger twin registers and also includes family members of twins, collecting data on twins as well as siblings, parents, spouses and offspring. The NTR conducts longitudinal survey and experimental studies with the help of registered twins and their family members. Information on young twins is obtained from parents and teacher reports, while adolescent and adult twins and their registered family members provide the data themselves[74]. Here, we compare adult twin individuals to their singleton siblings for a wide range of variables collected in survey and biobanking studies. To this aim, we selected at random one twin and a sibling from the same family, with the sibling of the same sex as the twin. We also selected only those siblings of twins who did not differ more than 6 years in age from the twin. For more detailed information on the methodology, please see the **Supplement**. The focus of most previous NTR studies has been on the quantification of genetic and non-genetic influences for a wide domain of traits, and while siblings were often included in the study design, relatively few studies reported the outcomes of twin-sibling comparisons. We identified the NTR studies which explicitly tested and reported the outcomes of twin-sibling comparisons and added these results to our discussion of twin-sibling differences in traits from various domains.

The Outcomes of Twin-Sibling Comparisons Across Multiple Domains in the NTR

Tables 1 and **2** summarize the findings for the twin-sibling comparisons for survey and biobank data, respectively.

Body Composition and Physical Development

Our findings for body composition as presented in **Table 1** and **Table 2** show that adult twins differed significantly from their singleton siblings in height and body mass index, with twins being somewhat smaller and lighter than their non-twin brothers and sisters. This was seen when data were obtained in the survey and trend was present for the data collected during the home visit, when weight was measured. In line with the trend for lower body mass index, twins also tended to have a smaller waist circumference at the time of the home visit.

Our results are in line with another large study in childhood and adolescence in the NTR, where twins were shown to be shorter and have a lower BMI than their siblings[75]. In a subset of this sample, they found the expected the twin-sibling difference in birth length and birthweight, with the effect still present at the age of 1 year but found no evidence for twin-sibling difference for height, weight and BMI at age 4. For this sample, no significant twin-sibling differences in body composition were seen any more at the young adult age, though there was a trend for twins to have a somewhat lower weight and shorter leg length. Additional components of physical development were also examined. While differences in growth hormone levels were seen, with on average lower levels of DHEAS and IGF-I levels in twins compared to their siblings[76], male twins did not differ from their siblings in testis size[77] and female twins did not differ for age at menarche[78].

These results and those of previous studies indicate that twins, who are more often born after a shorter gestation period and weigh less at birth than their singleton siblings, remain somewhat shorter and lighter well into adulthood, but in other aspects develop in the same way as their siblings.

Personality and Psychopathology

Table 1 presents the data for our matched twin-sibling comparison for five personality traits. Few differences are apparent, with only a trend for a personality trait called “openness to experience”, where twin seems to score somewhat lower than the siblings. For sensation seeking traits twins and siblings are similar, and twins also did not differ from siblings in their perception of support, life satisfaction, or loneliness. Several NTR studies previously tested and reported on adult twin-sibling differences in personality, in large samples that did not employ within-family designs, and reported no twin-singleton differences for the traits studied, which included sensation seeking[79], neuroticism[80] or trait anger[81].

With respect to psychopathology, **Table 1** shows twins did not differ from their siblings in ADHD symptoms in adults. They also did not differin borderline personality total scores, confirming previous findings in an overlapping sample[81], nor for anxious depression. Depression, anxiety and a combined anxious depression score were the subject of several previous studies in large samples of adult twins and singleton siblings, but no matter the definition, the two groups did not differ[13,14,82]. Adult twins and singleton siblings were also similar in the prevalence of burnout[13,15]. Earlier NTR studies did not show evidence for differences between twins and siblings for obsessive-compulsive symptoms in adults83 nor, in sample of young adults, for autistics traits[84]. A comparison of younger twins and siblings with respect to psychopathology showed no differences in ADHD symptoms between adolescent twins and singleton siblings[85]. **Table 1** also shows twins

TABLE 1.
Outcomes of within-family twin-sibling comparisons for body size, personality, mental health, demographics and lifestyle data collected in adult participants (survey 8) from the Netherlands Twin Register
Twins and siblings from the same family were selected to be of the same sex and of similar ages (not more than 6 years apart in age).

	Pairs	Twin		Sibling		Twin-sibling comparison results	
	N	Mean	SD	Mean	SD	T-test value	*p*-value
Continuous traits							
Age (years)	685	30.41	12.86	32.14	12.20	-15.347	0.000
Height (cm)	655	173.73	8.93	174.40	8.50	-2.480	0.013
Body Mass Index (BMI, kg/m^2)	634	22.73	3.12	23.89	3.82	-7.365	0.000
NEO neuroticism score	613	29.38	7.60	29.90	7.29	-1.400	0.162
NEO extraversion score	613	42.97	5.94	42.50	6.00	1.510	0.132
NEO openness to experience score	613	36.59	5.78	37.37	5.82	-2.734	0.006
NEO agreeable-ness score	613	45.54	4.97	45.01	4.84	2.206	0.028
NEO conscien-tiousness score	613	44.76	5.78	45.04	5.54	-0.948	0.344
SSS total sensation seeking score	411	11.22	2.41	11.13	2.45	0.711	0.478
SSS thrill adventure seeking score	605	9.16	3.65	8.83	3.69	1.942	0.053
SSS experience seeking score	418	16.00	4.46	16.33	4.58	-1.205	0.229
SSS boredom susceptibility score	593	13.45	4.79	18.30	4.82	-0.303	0.762
SSS disinhibition score	593	23.28	3.70	13.18	3.52	1.566	0.118
UNC-FSSQ confidant support	561	561	2.81	22.98	3.06	1.782	0.075
UNC-FSSQ affective support	559	13.81	1.97	13.72	1.99	0.745	0.457
SWLS general satisfaction with life score	649	27.67	4.91	27.34	5.17	1.299	0.194
TILS loneliness score	633	3.90	1.20	3.96	1.26	-0.920	0.358
CAARS ADHD index	599	8.11	4.12	8.44	3.93	-1.554	0.121
PAI-BOR total borderline personality score	603	15.16	8.37	15.92	8.22	-1.775	0.076
PAI-BOR affect instability	603	4.40	3.00	4.59	3.00	-1.225	0.221
PAI-BOR identity problems	603	3.82	2.72	4.05	2.90	-1.562	0.119
PAI-BOR negative relationships	602	4.14	2.85	4.62	2.75	-3.244	0.001
PAI- BOR self-harm	602	2.81	2.40	2.65	2.31	1.242	0.215
ASR anxious-depressed scale	526	4.93	5.15	5.29	5.46	-1.183	0.237

Table 1. (continued)

	Pairs	Twin		Sibling		Test statistics	
Dichotomous traits	***N***	***N* yes**	**% yes**	***N* yes**	**% yes**	**Chi-square**	***p*-value**
Being in good subjective health	657	590	89.8	575	87.5	1.675	0.196
Ever been in contact with mental health services	532	128	24.1	160	30.1	5.339	0.021
Regular sport participation	634	415	65.5	399	62.9	0.945	0.331
Regular alcohol use (2 or more times per week)	618	233	37.7	248	40.1	0.912	0.340
Current smoker	613	217	35.4	203	33.1	0.929	0.335
Ever tried hash	591	167	28.3	178	30.1	0.578	0.447
Being in a steady relationship, when 30 years or older	312	275	88.2	280	89.7	0.271	0.603
Living together with partner, when 30 years or older	312	258	82.7	268	85.9	1.095	0.295

Note: Abbreviations: NEO, NEO Five-Factor Inventory; SSS, Sensation Seeking Scale; UNC- FSSQ, Duke-UNC Functional Social Support Questionnaire; SWLF, Satisfaction With Life Scale, TILS, Three-Item Loneliness Scale; CAARS, Conners' Adult ADHD Rating Scales; PAI-BOR, Personality Assessment Inventory-Borderline Features scale; ASR, Adult Self-Report.

and singleton siblings were similar in their reports of being in good health and in ever having been in contact with mental health services.

Considering the overall picture of these findings, twins do not seem to differ from their siblings in personality and psychopathology. Any differences found were very small.

Behavioral and Sociodemographic Traits

Table 1 also presents the results of the twin-sibling comparison for various health behaviors, which show twins to be similar to their singleton siblings in their reports of regularly sport participation, regular alcohol drinking, current smoking behavior, and ever having tried hash. This is in line with NTR studies, in which the prevalence of problem drinking[86] and cannabis use initiation[87] was similar in twins and their siblings. Likewise, previous NTR studies that examined aspects of childhood behavior have not found differences between twins and siblings for bullying and victimization in 9-year-old twins and siblings[88] and for truancy during secondary education[89].

With respect to sociodemographic traits, **Table 1** indicates that there are no differences between twins and siblings for being in a relationship, and for living together with a partner. As age may play a role, we limited our comparison for relationship status to those aged 30 or older. A previous study by Middeldorp et al.[13] showed that at the age of 27 years monozygotic female twins were less often in a relationship than siblings. NTR studies on other sociodemographic traits did not find any twin-sibling differences for employment status[90] and the prevalence of being in a creative profession[91]. However, compared to their

singleton siblings, twins more often lived in highly urbanized areas in two younger age cohorts, though no difference was seen for the oldest age cohort[92]. It is possible that in the younger cohorts the age difference between twins and siblings was of importance, as the on average older sibling may already have had the financial means or the motive of an own increasing family to move to less populated areas.

Overall, twin status does not affect choices related to health behavior and most other behaviors, though particular aspects of social behavior such as being in a relationship and residential choices deserve more attention to provide definite answers.

Biomarkers and Disease

The results for our twin-sibling comparisons for biomarkers can be seen in **Table 2**. For total cholesterol, HDL cholesterol, LDL cholesterol, and triglyceride levels, there were no significant differences between twins and siblings. With respect to glucose metabolism, looking at the data the pattern seems to suggest a trend for siblings to have higher values, in particular for insulin. However, differences in BMI between twins and siblings are strongly related to differences in glucose metabolism parameters and when comparing scores after the effects of BMI are taken into account, twins and siblings were similar in glucose metabolism. We further compared twins and siblings on C-reactive protein, white and red blood cell count and IL6 level and found no significant differences for any of these variables. This is in line with a previous small-scale study including 222 twins and 85 siblings in which no differences between twins and siblings were observed for the cytokine response to *ex vivo* amyloid-beta stimulation[93]. **Table 2** also summarizes the twin-sibling data for liver enzymes and creatinine, which were similar for the two groups. This extends the findings by van Beek at al.[94] , who in a larger sample that included this subset, also reported no differences in the liver enzymes aspartate aminotransferase (AST) and alanine aminotransferase (ALT) but did not comment on gamma-glutamyl transferase (GGT).

While the majority of our comparisons did not reveal any twin-sibling differences, there was a trend for twin and siblings to differ in telomere length, with shorter telomere length in the siblings compared to the twins, even when correcting for age effects using standardized scores. Further studies would be needed to determine whether this represents a meaningful twin-sibling difference.

A series of previous NTR studies focused on cardiovascular functioning in adult twins and their siblings. No differences were seen for blood pressure, whether measured in a laboratory setting[95], ambulatory during everyday life[96], or defined as hypertensive state[97]. For cardiac functioning, operationalized as heart rate or heart rate variability, again no significant differences between twins and their siblings were found[98,99]. Twin-sibling differences were also not present for respiration rate[99] or cortisol[100].

With respect to other physical diseases, NTR studies in adult women showed twins did not differ from their singleton siblings in the occurrence of polycystic ovary syndrome based on survey information[101], nor in the presence of cervix smear abnormalities as determined by cervical screening[102]. Twin-sibling differences were also not reported for asthma in a study including more than 11,000 adult male and female twins and siblings, but surprisingly siblings reported somewhat higher rates of allergy than twins (23.3% vs 18.2%[103]).

In summary, these studies show that, while there may be a few exceptions pointing to a disadvantage in singletons, twins and their singleton siblings are generally very similar in biomarkers and disease.

Cognitive Function

We did not a apply within-family test for cognitive function in the present study, but here summarize the results from previous NTR studies. A number of studies in young twins have shown differences in cognitive abilities. In samples of twins and siblings with average age between 9 and 13, twins scored lower on full scale, verbal and performance IQ[104] and on IQ, reading performance and verbal working memory[105]. De Zeeuw et al.[106] tested cognitive function in the largest sample yet, including 1375 twins and 1375 siblings and again showed lower scores for specific aspects of cognitive function (IQ, reading performance and verbal working memory) at age 9 in twins compared to their siblings. However, twins are often not the first-born children in a family. After taking birth order into account the cognitive disadvantage of twins dissipated. Lower scores for cognitive function was only seen when comparing the twins to older siblings. When comparing the scores of twins at age 9 with the scores of their younger siblings at age 9, twins no longer differed from their singleton siblings. Twins did not differ from older or younger siblings with respect to for visuospatial working memory, verbal and spatial short-term memory. Interestingly, twins scored higher on physical education than their older and younger singleton siblings[106]. Note that not all studies on cognition in young twins showed lower scores in twins compared to their singleton siblings. A recent study including more than 11,000 twins and 262 of their siblings at age 7.5 that reading ability in twins was comparable to that in siblings and to national norms[107]. Two smaller NTR studies provided information on the twin-sibling comparison for adult cognitive function: Van den Berg et al.[108] found no twin-sibling differences for reading vocabulary, and Posthuma et al.[109] showed no difference in intelligence between twins and siblings.

Overall, while lower scores for cognitive ability in twins have reported, these effects seem related to the position (birth order) in the nuclear family. Those first born in the family more often score higher on cognitive tests than those later born, an effect which is independent of whether they are born from a singleton or multiple pregnancy. As not all studies in children showed lower scores across the cognitive domains and studies in adults did not show any twin-singleton differences, a twin-singleton difference for cognitive function seems modest and likely limited to specific domains.

Concluding Remarks

A twin or multiple pregnancy carries an increased risk of prenatal mortality and congenital abnormalities. However, the majority of children born from a multiple pregnancy are healthy at birth and develop into healthy individuals. Still, multiples could differ from singletons in other ways, due to their lower gestational age and birthweight or due to growing up with a sibling of the same age and often same sex and looks, i.e. the close companionship hypothesis. Our studies indicate that twins do not differ from their singleton siblings across a wide range of behavioral and lifestyle parameters, biomarkers, or diseases.

TABLE 2.
Within-family twin-sibling comparison for biomarkers as assessed in the Netherlands Twin register
Twins and siblings from the same family were selected to be of the same sex and of similar ages (not more than 6 years apart in age).

	Pairs	Twin		Sibling		Correlation of trait difference with		Twin-Sibling comparison results using standardised scores[a]	
	N	Mean	SD	Mean	SD	Difference in age	Difference in BMI	T-test value	*p*-value
Age (years)	382	36.30	12.28	37.37	11.98	1.00	.084	-6.586	0.000
Height (cm)	378	174.37	9.02	175.58	9.13	-.041	-.089	-3.855	0.000
Body mass index (BMI)	374	24.37	4.05	25.29	4.37	.084	1.00	-3.399	0.001
Waist circumference (cm)	373	83.30	11.26	85.72	12.23	.062	.809***	-3.153	0.002
Fasting total cholesterol (mmol/l)	323	4.85	1.02	4.94	.94	.103	.074	-.911	0.363
Fasting HDL (mmol/l)	323	1.40	.35	1.41	.37	.053	-.258***	-2.118	0.035
Fasting LDL (mmol/l)	323	2.93	.92	2.95	.84	.089	.096	-.223	0.824
Triglycerides (mmol/l)	323	1.16	.57	1.27	.63	.033	.277***	-1.450	0.148
Fasting Glucose (mmol/l)	321	5.30	.73	5.41	.88	-.000	.266***	-1.140	0.255
Fasting Insulin (μIU/ml	315	8.41	4.83	9.82	7.94	.055	.392***	-1.539	0.125
Hba1c (%)	318	5.26	.51	5.33	.64	.116	.051	-1.388	0.166
White blood cell count (10^{12}/L)	346	6.33	1.72	6.58	1.94	.020	.165**	-1.690	0.092
Red blood cell count (10^{12}/L)	347	4.68	.48	4.66	.46	.022	.164**	1.470	0.143
C-reactive protein (mg/L)	353	2.69	4.18	3.47	5.00	.030	.299***	-1.726	0.085
Interleukin 6 (pg/mL)	347	1.42	1.27	1.97	5.24	.093	.043	-1.758	0.080
AST (U/L)	293	20.83	6.92	20.68	7.22	.075	.113	.701	0.484
ALT (U/L)	275	10.93	7.39	11.02	9.19	.085	.072	.314	0.754
GGT (U/L)	293	28.03	24.47	29.87	24.00	.081	.184***	-.399	0.690
Creatinine (U/L)	293	85.52	13.95	85.28	14.26	-.058	.046	.420	0.674
Telomere length	289	2.78	.69	2.66	.49	-.209***	-.045	3.491	0.001

Note: *Abbreviations:* ALT, enzymes alanine transaminase, AST, aspartate transaminase, GGT, gamma-glutamyltransferase
* < 0.05, ** < 0.01, ***< 0.001
a Age and sex standardised scores were used in the paired sample t-tests for height, BMI and waist circumference, for all other variables age, sex and BMI standardised scores were used.

The one aspect on which twins differed from non-twins was body composition. Twins remained smaller and lighter compared to their singleton siblings, even as adults, attaining about 1 point lower BMI lower than singletons. As very few physiological differences were seen, the difference in body composition seems to be a lasting and likely beneficial aspect of being a twin.

Very little evidence was found for the close companionship hypothesis: twins did not differ from their singleton siblings across a wide range of behavioral and psychological traits. Twins did do better than singletons in physical education classes at school, which could indicate the effect of always having a playmate during childhood. In adults, no differences in regular sport participation are seen, suggesting that the effect, if confirmed, may only be present at younger ages. We did not test for effects of zygosity and we cannot exclude the possibility that specific twin effects may still occur for specific groups, especially monozygotic twins, as implied by the Middeldorp et al.[13] finding that monozygotic female twins were less often in a relationship than others. However, as our study included a large number of monozygotic twins the effects are likely to be small and limited to specific situations and groups.

We did not investigate effects of fertility treatment. Many of the studies presented here were conducted in twins who were born before ART and IVF became a frequently used in the Netherlands, but it may have occurred in the younger twins participating in the studies. Still, many of these twins will not have had a sibling, as fertility problems occurred in these families, and the number of IVF twin-sibling pairs will thus be very limited. A previous study by the NTR using matching of IVF with DZ naturally conceived twins showed no differences in growth, attainment of motor milestones, or in behavioral development, leading to the conclusion that for nearly all aspects, development in these groups of children is similar[110].

While twins and siblings born in the same family do not differ, it is still possible that individuals born in twin families differ from singletons born within non-twin families due to the genetic background of their parents. This may be especially so in dizygotic twin families, as we know that mothers of DZ twins are somewhat taller and heavier than mothers of monozygotic twins[111], and genes found to be related to twinning are also related to increased body mass[71]. In a study of 5-year-old Dutch twin children, female twins were as tall as singleton children, while male twins were still somewhat shorter than children from the general population and twins overall had a lower BMI than the general population[112]. In a separate study, Estourgie-van Burk et al.[75] reported no differences in height between young adult twins and their siblings and their height was comparable to the general population. For BMI, no differences were observed between 18-year-old twins and 18-year olds from the general population, whereas the siblings of twins had increased BMI values. A Finnish study among 17-year-old twins reported that twins were as tall as singletons, but that boy twins were still leaner, though an American twin study showed no twin effects on weight and height at age eight[113]. It is possible that, when it comes to body weight and body height, twins on average reach their full potential at a later age than singletons. Alternatively, twins may never reach their full potential for height and weight. As a large number of the genetic variants involved in body composition are now known, this hypothesis could be tested in the near future. Other factors may also differ

within families for twins and non-twins. Maternal age at birth and parity are higher on average for dizygotic twins, and twins differ from singletons in gestational age. Also, the rearing environment is different for two children than for a single child, though this is not specifically limited to multiples but is the case for any family including multiple children.

Another aspect that may deserve additional attention is the position of multiples in the family. For a number of cognitive traits, the differences between twins and siblings could be explained by position within the family. This has not been systematically studied for other traits, which may reveal similar position-in-family effects. In addition to the position in the family, whether twins were born first or second, also requires more attention. In this study, there was an equal distribution of twins born first and second by design. As birth complications more often occur for the second born twin, this may be another factor leading to small differences between twins and singletons.

Our results regarding a large number traits and common disorders are fairly optimistic, indicating that twins do not differ from singletons. By comparing the twins to their non-twin siblings, we avoid confounding by between-family factors. We note, however, that our work did not look at rare disorders and that despite the large number of participants in the Netherlands Twin Register some forms of bias may be present. Parents of twins who presented with serious complications at birth may decide against participation. Adolescent and adult twins with health problems, whether mental or physical, may be less inclined to enroll in the longitudinal study or may drop out during the study. In addition, the decision to enroll and continue participation is also influenced by other factors such as educational attainment. Still, twins are born in all strata of society and tend to be motivated to take part in medical and scientific studies and, though we have seen that NTR participation is related to educational attainment, participants with lower educational levels are also present in the sample.

In conclusion, with the exception of congenital disorders and body composition, twins do not seem to differ from singletons, when taking family factors into account. While being a twin can be seen as special, for most traits twins are just like ordinary siblings.

Quotes by twins

In the ninth NTR survey to adult participants[74], twins were asked whether they liked being a twins. The following quotes are a selection from answers provided by more than 5000 multiples. Their comments are in line with the findings in our chapter.

They highlight that a multiple pregnancy carries risk:

• “I did not grow up as twins. Twin sister died at birth”

Many of the multiples view their multiple status as something special, though not everyone responds positively.

• “I like it as long as people see you as separate individuals. It is in any case not ordinary and I do like that”

• “Sad for all those singletons…honestly”

• “Quite nice. I have different (closer) contact with my twin brother than with my other brothers”

• “Quite nice; as twin you are never alone”

• “Being a twin has two sides. A very nice side because we got along very well as sisters and because being a twin is also somewhat special. The disadvantage of being a twin is that I have the feeling I am constantly compared to my twin sisters and that I do this also.”

• “Still somewhat special. You do have another connection with each other than with my other brothers (4)”

• “Super great! Feeling that there is someone who always has my back and who understand me. This is I think because we are also friends”

• “In general nice, because we understand each other very well and are strongly connected. But sometimes it is also a bit suffocating since you were and are always compared to each other and we also compare ourselves and never want to do under for the other”

• “Nice. You have a special connection with each other which you do not have with brother or sister. A bit difficult to learn to make friends and keep them, because I always had my sister and never had to do this”

However, comments also indicate that the multiples do not see much difference with ordinary brothers and sisters, especially not as adults.

• “We are a dizygotic twin pair, so I do not see it as very special. I do not connect more with him than with my other brothers”

• “We have a very different life. I view him just like my other brother”

• “It’s OK. Not very special. The connection with my younger sister is no less, my younger sister and my twin sister are just as important for me. The one is not more important than the other. Sometimes it is nice to be a twin, sometimes less nice. It has (in the past) advantages and disadvantages”

• “Quite nice, not very special”

• “Quite nice but not as special as people often expect”

• “It is nice that you can race to see who finishes his study first, but noting more”

• “It is somewhat more special than other brothers/sister, but for the rest not much difference”

• “In the past special to do everything together, now I do not notice much. It is just like an ordinary brother”

• “In the past very nice, now it just seems as if we are normal sisters”

• “OK. In the past I did a lot together with my twin brothers Now it is just like with my other brother or sister”

• “Normal. No differences with people who are not a twin”

• “Not different than having my other two brothers”

Supplement Within-Family Twin – Non-Twin Comparisons Across a Wide Range of Traits Assessed in the Netherlands Twin Register

Participants and selection procedure

To compare twins to their siblings, we made use of the data from two separate projects: (1) the eighth NTR survey on health, lifestyle and personality, which was sent out to adult participants between 2004 and 2009[74,114]; (2) a large-scale blood collection project carried out between 2004 and 2009 in which participants were visited at home to obtain blood samples and health information[115].

For each of the two datasets, we followed the same procedure. We first selected all twins and their singleton siblings with known age, sex, birth year, and, in the case of twins, zygosity. Known half-siblings and non-biological siblings were excluded. Next, we randomly selected one of the twins in case a twin pair both participated and assessed whether there was a same-sex singleton sibling in the family who was born within 6 years from the twin. This sibling was then selected, if at the time of participating he or she was within 6 years of the age of the twin. In the case multiple singleton siblings of a twin met the criteria, we selected the same-sex sibling closest in birth year. We made two exceptions to this procedure. In the case of an opposite-sex twin pair, when a male twin-sibling pair could be formed in the dataset, we selected the male pair. This was done to maximize the presence of male pairs in the analyses as fewer males participated than females. When both older and younger siblings within the 6-year time frame were present, preference was given to the sibling younger than the twin, as twins more often have older siblings than younger siblings. Even applying these criteria, our sample selections included more female pairs than male pairs and more older than younger singleton siblings.

NTR Survey 8

As part of a longitudinal survey study, this survey was sent out to adult twins registered with the NTR. It was completed by 10,176 multiples and 2,142 siblings and collected information on a wide range of traits[114]. After applying the selection criteria, the sample for the present analyses consisted of 685 twin-sibling pairs, 177 (26%) being male. Average age of the twins at the time of survey completion was 30.4 (SD=12.9) and of the siblings 32.1 (SD=12.2) years. In 142 (21%) of the pairs, the twin was older than the sibling. We compared the twins and their singleton siblings on the following continuous traits: self-reported height and body mass index (BMI, calculated as $weight_{(kg)}/height^2_{(m)}$); the big five personality dimensions (openness, conscientiousness, extraversion, agreeableness, neuroticism) as measured with the NEO Five-Factor Inventory[116,117]; borderline personality components measured with the Personality Assessment Inventory-Borderline Features scale (PAI-BOR118); sensation seeking score and its subscales thrill and adventure

seeking, boredom susceptibility, disinhibition and experience seeking as measured with the Sensation Seeking Scale[119,120]; Attention deficit hyperactivity disorder (ADHD) as measured with the Conners' Adult ADHD Rating Scales (CAARS[121]); the anxious depression scale of the Adult Self-Report which combines elements of depression and anxiety[122]; dimensions of social support (confidant and affective) as measured with the Duke-UNC Functional Social Support Questionnaire; and life satisfaction as measured with the Satisfaction With Life Scale[123]; loneliness as measured with the Three-Item Loneliness Scale[124]. In addition, we compared twins and siblings on categorical traits with outcomes operationalized as yes versus no, including the following traits: being in good health (reports of good or excellent health were coded as yes), ever been in contact with mental health services, being a current smoker, drinking alcohol regularly (reports of drinking alcohol 2 or more times per week were classified as yes), ever tried hash, being in a relationship (data only included for those age 30 years and older) and living together with a partner (data only included for those aged 30 years and older).

NTR Biobank

Between 2004 and 2010, 9530 individuals provided a blood sample and health-related information as part of a large-scale biobank project[115]. When conducting the selection procedure as described above, this resulted in 382 twin-sibling pairs, of which 144 (38%) were male-male pairs. The average age of the twins was 36.3 (12.3) and of the siblings 37.4 (12.0). In 265 pairs the singleton sibling was older than the twin. We compared twins and their singleton siblings on the following variables: lipid profile (total cholesterol, HDL, LDL and triglyceride levels), glucose metabolism (glucose, insulin and HbA1c levels), white and red blood cell counts, C-reactive protein as indicator of general inflammation, liver enzymes alanine transaminase (ALT), aspartate transaminase (AST) and gamma-glutamyltransferase (GGT), creatinine as measure of kidney function and telomere length. For lipid profile and glucose metabolism, data were only included if the participant had kept to the instruction to be fasting at the time of blood collection[115].

Analyses

All analyses were conducted in IBM SPSS Statistics version 25. To compare the twins with their singleton sibling we conducted a paired-sample t-test for continuous traits and a McNemar chi-squared test for categorical traits. As age and BMI may be important factors in the physiological parameters, we here correlated the differences in age and BMI for the twins and siblings with their differences in physiological parameters and present the test outcomes for age, sex and BMI standardized residuals. Considering the large number of comparisons conducted, we consider a trend when *p*-values are between 0.010 and 0.001 and p-values < 0.001 as significant.

References

1. Boyle, B. B., Helen, D. & McConkey, R. The Epidemiology of Congenital Anomalies in Multiple Births: A European Registry Based Study. *Pediatric Research* **70**, 408–408 (2011).
2. Zeitlin, J. *et al.* The European Perinatal Health Report: comparing the health and care of pregnant women and newborn babies in Europe. *Journal of Epidemiology & Community Health* **63**, 681–682 (2009).
3. Pison, G., Monden, C. & Smits, J. Twinning Rates in Developed Countries: Trends and Explanations. *Population and Development Review* **41**, 629–649 (2015).
4. Lambalk, C. B., de Koning, C. H. & Braat, D. D. The endocrinology of dizygotic twinning in the human. *Molecular and cellular endocrinology* **145**, 97–102 (1998).
5. Gielen, M. *et al.* Secular trends in gestational age and birthweight in twins. *Human reproduction (Oxford, England)* **25**, 2346–2353 (2010).
6. Mook-Kanamori, D. O. *et al.* Heritability estimates of body size in fetal life and early childhood. *PloS one* **7**, e39901 (2012).
7. Barker, D. J. In utero programming of chronic disease. *Clinical science (London, England : 1979)* **95**, 115–128 (1998).
8. Kwon, E. J. & Kim, Y. J. What is fetal programming?: a lifetime health is under the control of in utero health. *Obstetrics & gynecology science* **60**, 506–519 (2017).
9. Kajantie, E. *et al.* Size at birth as a predictor of mortality in adulthood: a follow-up of 350 000 person-years. *International journal of epidemiology* **34**, 655–663 (2005).
10. Carey, G. Twin imitation for antisocial behavior: implications for genetic and family environment research. *Journal of abnormal psychology* **101**, 18–25 (1992).
11. Veldkamp, S. A. M. *et al.* Parental Age in Relation to Offspring's Neurodevelopment. *Journal of Clinical Child & Adolescent Psychology* **50**, 632–644 (2021).
12. Zondervan-Zwijnenburg, M. A. J. *et al.* Parental Age and Offspring Childhood Mental Health: A Multi-Cohort, Population-Based Investigation. *Child development* **91**, 964–982 (2020).
13. Middeldorp, C. M., Cath, D. C., Vink, J. M. & Boomsma, D. I. Twin and genetic effects on life events. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **8**, 224–231 (2005).
14. Middeldorp, C. M. *et al.* Familial clustering of major depression and anxiety disorders in Australian and Dutch twins and siblings. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **8**, 609–615 (2005).
15. Middeldorp, C. M., Stubbe, J. H., Cath, D. C. & Boomsma, D. I. Familial clustering in burnout: a twin-family study. *Psychological medicine* **35**, 113–120 (2005).
16. Boyle, B. *et al.* Trends in the prevalence, risk and pregnancy outcome of multiple births with congenital anomaly: a registry-based study in 14 European countries 1984-2007. *BJOG: An International Journal of Obstetrics & Gynaecology* **120**, 707–716 (2013).
17. Doom, E. C. G., Delbaere, I., Martens, G. & Temmerman, M. Birth weight for gestational age among Flemish twin population. *Facts, views & vision in ObGyn* **4**, 42–49 (2012).
18. Cunningham, F. G. *et al.* Fetal-Growth Disorders. in *Williams Obstetrics, 25e* (McGraw-Hill Education, 2018).
19. Muhlhausler, B. S., Hancock, S. N., Bloomfield, F. H. & Harding, R. Are Twins Growth Restricted? *Pediatric Research* **70**, 117–122 (2011).

20. Groene, S. G., Tollenaar, L. S. A., Oepkes, D., Lopriore, E. & van Klink, J. M. M. The Impact of Selective Fetal Growth Restriction or Birth Weight Discordance on Long-Term Neurodevelopment in Monochorionic Twins: A Systematic Literature Review. *Journal of clinical medicine* **8**, (2019).

21. Sheay, W., Ananth, C. v. & Kinzler, W. L. Perinatal Mortality in First- and Second-Born Twins in the United States. Obstetrics & Gynecology 103, 63–70 (2004).

22. Vasak, B. *et al.* Lower perinatal mortality in preterm born twins than in singletons: a nationwide study from The Netherlands. *American Journal of Obstetrics and Gynecology* **216**, 161.e1-161.e9 (2017).

23. Doyle, P. The outcome of multiple pregnancy. *Human reproduction (Oxford, England)* **11** Suppl 4, 110–120 (1996).

24. Heino, A. *et al.* Variations in Multiple Birth Rates and Impact on Perinatal Outcomes in Europe. *PloS one* **11**, e0149252 (2016).

25. Kahn, B. *et al.* Prospective risk of fetal death in singleton, twin, and triplet gestations: implications for practice. *Obstetrics & Gynecology* **102**, 685–692 (2003).

26. Alexander, G. R., Slay Wingate, M., Salihu, H. & Kirby, R. S. Fetal and Neonatal Mortality Risks of Multiple Births. *Obstetrics and Gynecology Clinics of North America* **32**, 1–16 (2005).

27. Hansen, M. *et al.* Twins born following assisted reproductive technology: perinatal outcome and admission to hospital. *Human Reproduction* **24**, 2321–2331 (2009).

28. Hofmeyr, G. J., Barrett, J. F. & Crowther, C. A. Planned caesarean section for women with a twin pregnancy. *The Cochrane database of systematic reviews* CD006553 (2011) doi:10.1002/14651858.CD006553.pub2.

29. Rossi, A., Mullin, P. & Chmait, R. Neonatal outcomes of twins according to birth order, presentation and mode of delivery: a systematic review and meta-analysis*. *BJOG: An International Journal of Obstetrics & Gynaecology* **118**, 523–532 (2011).

30. Svenvik, M., Brudin, L. & Blomberg, M. Preterm Birth: A Prominent Risk Factor for Low Apgar Scores. *BioMed Research International* **2015**, 1–8 (2015).

31. Thorngren-Jerneck, K. & Herbst, A. Low 5-minute Apgar score: a population-based register study of 1 million term births. *Obstetrics and gynecology* **98**, 65–70 (2001).

32. Hartley, R. S. & Hitti, J. Birth order and delivery interval: Analysis of twin pair perinatal outcomes. *The Journal of Maternal-Fetal & Neonatal Medicine* **17**, 375–380 (2005).

33. Odintsova, V. v *et al.* Pre- and Perinatal Characteristics Associated with Apgar Scores in a Review and in a New Study of Dutch Twins. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **22**, 164–176 (2019).

34. Scher, A. I. *et al.* The risk of mortality or cerebral palsy in twins: a collaborative population-based study. *Pediatric research* **52**, 671–681 (2002).

35. Pharoah, P. O. D. Risk of Cerebral Palsy in Multiple Pregnancies. *Clinics in Perinatology* **33**, 301–313 (2006).

36. Pharoah, P. O. D. & Dundar, Y. Monozygotic twinning, cerebral palsy and congenital anomalies. *Human reproduction update* **15**, 639–648 (2009).

37. Bonellie, S., Currie, D. & Chalmers, J. Comparison of risk factors for cerebral palsy in twins and singletons. *Developmental Medicine & Child Neurology* **47**, 587–591 (2005).

38. Glinianaia, S. V. v, Rankin, J. & Wright, C. Congenital anomalies in twins: a register-based study. *Human reproduction (Oxford, England)* **23**, 1306–1311 (2008).

39. Myrianthopoulos, N. C. Congenital malformations in twins. *Acta geneticae medicae et gemellologiae* **25**, 331–335 (1976).

40. Project, E.-P. & Euro-Peristat, P. European perinatal health report. Core indicators of the health and care of pregnant women and babies in Europe in 2015. (2018).

41. Hall, J. G. Twinning. *Lancet (London, England)* **362**, 735–743 (2003).

42. Källén, B. Congenital malformations in twins: a population study. *Acta geneticae medicae et gemellologiae* **35**, 167–178 (1986).

43. Li, S.-J., Ford, N., Meister, K. & Bodurtha, J. Increased risk of birth defects among children from multiple births. *Birth Defects Research Part A: Clinical and Molecular Teratology* **67**, 879–885 (2003).

44. Mastroiacovo, P. *et al.* Congenital malformations in twins: an international study. *American journal of medical genetics* **83**, 117–24 (1999).

45. Pradat, P. Epidemiology of major congenital heart defects in Sweden, 1981-1986. *Journal of epidemiology and community health* **46**, 211–215 (1992).

46. Windham, G. C. & Bjerkedal, T. Malformations in twins and their siblings, Norway, 1967-79. *Acta geneticae medicae et gemellologiae* **33**, 87–95 (1984).

47. Windham, G. C. & Sever, L. E. Neural tube defects among twin births. *American journal of human genetics* **34**, 988–98 (1982).

48. Yu, Y. *et al.* Birth Anomalies in Monozygotic and Dizygotic Twins: Results From the California Twin Registry. *Journal of epidemiology* **29**, 18–25 (2019).

49. Grosen, D. *et al.* Risk of oral clefts in twins. *Epidemiology (Cambridge, Mass.)* **22**, 313–9 (2011).

50. Doyle, P. E., Beral, V., Botting, B. & Wale, C. J. Congenital malformations in twins in England and Wales. *Journal of epidemiology and community health* **45**, 43–48 (1991).

51. Boyle, B. *et al.* Prevalence and risk of Down syndrome in monozygotic and dizygotic multiple pregnancies in Europe: implications for prenatal screening. *BJOG: An International Journal of Obstetrics & Gynaecology* **121**, 809–820 (2014).

52. Bliek, J. *et al.* Lessons from BWS twins: complex maternal and paternal hypomethylation and a common source of haematopoietic stem cells. *European journal of human genetics: EJHG* **17**, 1625–1634 (2009).

53. Melnick, M. & Myrianthopoulos, N. C. The effects of chorion type on normal and abnormal developmental variation in monozygous twins. *American Journal of Medical Genetics* **4**, 147–156 (1979).

54. Pharoah, P. O. D. Causal hypothesis for some congenital anomalies. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **8**, 543–550 (2005).

55. Bergh, T., Ericson, A., Hillensjö, T., Nygren, K. G. & Wennerholm, U. B. Deliveries and children born after in-vitro fertilisation in Sweden 1982-95: a retrospective cohort study. *Lancet (London, England)* **354**, 1579–1585 (1999).

56. Ooki, S. Maternal age and birth defects after the use of assisted reproductive technology in Japan, 2004-2010. *International journal of women's health* **5**, 65–77 (2013).

57. Bryan, E., Little, J. & Burn, J. Congenital anomalies in twins. *Bailliere's clinical obstetrics and gynaecology* **1**, 697–721 (1987).

58. Singh, S., Murphy, B. & O'Reilly, R. Epigenetic contributors to the discordance of monozygotic twins. *Clinical Genetics* **62**, 97–103 (2002).

59. Oates, N. A. *et al.* Increased DNA methylation at the AXIN1 gene in a monozygotic twin from a pair discordant for a caudal duplication anomaly. *American journal of human genetics* **79**, 155–162 (2006).

60. Lancaster P, Kucera J, Knudsen L, Botting B, Robert-Gnansia E, Goujard J, Elek C, Mastroiacovo P, Cocchi G, Borman B, Irgens L, Castilla E, Martinez-Frias M-L, E. A. Conjoined twins--an epidemiological study based on 312 cases. The International Clearinghouse for Birth Defects Monitoring Systems. *Acta geneticae medicae et gemellologiae* **40**, 325–35 (1991).

61. Meyers, C., Elias, S. & Arrabal, P. Congenital anomalies and pregnancy loss. in *Multiple pregnancy: epidemiology, gestation, and perinatal outcome* (eds. L.G., K. & E., P.) **73–78** (Parthenon Publishing Group, 1995).

62. Lopriore, E., Nagel, H. T. C., Vandenbussche, F. P. H. A. & Walther, F. J. Long-term neurodevelopmental outcome in twin-to-twin transfusion syndrome. *American journal of obstetrics and gynecology* **189**, 1314–9 (2003).

63. Hack, K. E. A. *et al.* Increased perinatal mortality and morbidity in monochorionic versus dichorionic twin pregnancies: clinical implications of a large Dutch cohort study. BJOG : an international journal of obstetrics and gynaecology 115, 58–67 (2008).

64. Ortibus, E. *et al.* The pregnancy and long-term neurodevelopmental outcome of monochorionic diamniotic twin gestations: a multicenter prospective cohort study from the first trimester onward. *American Journal of Obstetrics and Gynecology* **200**, 494.e1-494.e8 (2009).

65. Layton, W. M. Situs inversus in conjoined twins. *American Journal of Medical Genetics* **34**, 297–297 (1989).

66. Kean, B. H. Complete transposition of the viscera in both of one-egg twins. *Journal of Heredity* **33**, 217–221 (1942).

67. Bondavalli, W., Bondavalli, G., la Bella, G. & Nizzoli, A. [A case of dextrocardia with concordant situs viscerum inversus totalis in pairs of dizygotic twins]. *Acta geneticae medicae et gemellologiae* **8**, 201–214 (1959).

68. Gedda, L. *et al.* Situs viscerum specularis in monozygotic twins. *Acta geneticae medicae et gemellologiae* **33**, 81–85 (1984).

69. Agirbasli, M., Hamid, R., Jennings, H. S. & Tiller, G. E. Situs inversus with hypertrophic cardiomyopathy in identical twins. *American Journal of Medical Genetics* **91**, 327–330 (2000).

70. Segal, N. L. Situs Inversus Totalis in Twins: A Brief Review and a Life History / Twin Research: Twin Studies of Trisomy 21; Monozygotic Twin Concordance for Bilateral Coronoid Hyperplasia; Prenatal Hormonal Effects in Mixed-Sex Non-Human Primate Litters; Insurance M. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **20**, 90–95 (2017).

71. Mbarek, H. et al. Identification of Common Genetic Variants Influencing Spontaneous Dizygotic Twinning and Female Fertility. *American journal of human genetics* **98**, 898–908 (2016).

72. Dick, D. M., Johnson, J. K., Viken, R. J. & Rose, R. J. Testing between-family associations in within-family comparisons. *Psychological science* **11**, 409–413 (2000).

73. Posthuma, D. *et al.* Theory and practice in quantitative genetics. *Twin research: the official journal of the International Society for Twin Studies* **6**, 361–376 (2003).

74. Ligthart, L. *et al.* The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-Family Designs. *Twin Research and Human Genetics* **22**, 623–636 (2019).

75. Estourgie-van Burk, G. F., Bartels, M., Boomsma, D. I. & Delemarre-van de Waal, H. A. Body size of twins compared with siblings and the general population: from birth to late adolescence. *The Journal of pediatrics* **156**, 586–591 (2010).

76. Estourgie-van Burk, G. F., Bartels, M. & Boomsma, D. I. A twin-sibling study on early growth and hormone levels in adolescents. *Behavior genetics* **45**, 283–293 (2015).

77. Estourgie-van Burk, G. F., Bartels, M., Delemarre-van de Waal, H. A. & Boomsma, D. I. Heritability of testis size. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **12**, 351–355 (2009).

78. Van den Berg, S. M. & Boomsma, D. I. The familial clustering of age at menarche in extended twin families. *Behavior genetics* **37**, 661–667 (2007).

79. Stoel, R. D., de Geus, E. J. C. & Boomsma, D. I. Genetic analysis of sensation seeking with an extended twin design. *Behavior genetics* **36**, 229–237 (2006).

80. Willemsen, G. & Boomsma, D. I. Religious upbringing and neuroticism in Dutch twin families. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **10**, 327–333 (2007).

81. Distel, M. A. *et al.* The covariation of trait anger and borderline personality: a bivariate twin-siblings study. *J Abnorm Psychol* **121**, 458–466 (2012).

82. Smit, D. J. A., Posthuma, D., Boomsma, D. I. & de Geus, E. J. C. The relation between frontal EEG asymmetry and the risk for anxiety and depression. *Biological psychology* **74**, 26–33 (2007).

83. Van Grootheest, D. S., Cath, D. C., Beekman, A. T. & Boomsma, D. I. Genetic and environmental influences on obsessive-compulsive symptoms in adults: a population-based twin-family study. *Psychological medicine* **37**, 1635–1644 (2007).

84. Hoekstra, R. A., Bartels, M., Verweij, C. J. H. & Boomsma, D. I. Heritability of autistic traits in the general population. *Archives of pediatrics & adolescent medicine* **161**, 372–377 (2007).

85. Polderman, T. J. C., van Dongen, J. & Boomsma, D. I. The relation between ADHD symptoms and fine motor control: a genetic study. *Child neuropsychology: a journal on normal and abnormal development in childhood and adolescence* **17**, 138–150 (2011).

86. De Moor, M. H. M. *et al.* Heritability of problem drinking and the genetic overlap with personality in a general population sample. *Frontiers in genetics* **2**, 76 (2011).

87. Distel, M. A. *et al.* Age moderates non-genetic influences on the initiation of cannabis use: a twin-sibling study in Dutch adolescents and young adults. *Addiction (Abingdon, England)* **106**, 1658–1666 (2011).

88. Veldkamp, S. A. M. *et al.* Bullying and Victimization: The Effect of Close Companionship. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **20**, 19–27 (2017).

89. Van der Aa, N., Rebollo-Mesa, I., Willemsen, G., Boomsma, D. I. & Bartels, M. Frequency of truancy at high school: evidence for genetic and twin specific shared environmental influences. *The Journal of adolescent health: official publication of the Society for Adolescent Medicine* **45**, 579–586 (2009).

90. Middeldorp, C. M., Cath, D. C. & Boomsma, D. I. A twin-family study of the association between employment, burnout and anxious depression. *Journal of affective disorders* **90**, 163–169 (2006).

91. Roeling, M. P., Willemsen, G. & Boomsma, D. I. Heritability of Working in a Creative Profession. *Behavior genetics* **47**, 298–304 (2017).

92. Willemsen, G., Posthuma, D. & Boomsma, D. I. Environmental factors determine where the Dutch live: results from the Netherlands twin register. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **8**, 312–317 (2005).

93. Posthuma, D. *et al.* Human cytokine response to ex vivo amyloid-beta stimulation is mediated by genetic factors. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **8**, 132–137 (2005).

94. Van Beek, J. H. D. A. *et al.* The genetic architecture of liver enzyme levels: GGT, ALT and AST. *Behavior genetics* **43**, 329–339 (2013).

95. Hottenga, J.-J. *et al.* Heritability and stability of resting blood pressure. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **8**, 499–508 (2005).

96. Kupper, N. *et al.* Heritability of daytime ambulatory blood pressure in an extended twin design. *Hypertension (Dallas, Tex.: 1979)* **45**, 80–85 (2005).

97. De Geus, E. J., Posthuma, D., Ijzerman, R. G. & Boomsma, D. I. Comparing blood pressure of twins and their singleton siblings: being a twin does not affect adult blood pressure. *Twin research: the official journal of the International Society for Twin Studies* **4**, 385–391 (2001).

98. Kupper, N. H. M. *et al.* Heritability of ambulatory heart rate variability. *Circulation* **110**, 2792–2796 (2004).

99. Kupper, N. *et al.* A genetic analysis of ambulatory cardiorespiratory coupling. *Psychophysiology* **42**, 202–212 (2005).

100. Kupper, N. *et al.* Familial influences on basal salivary cortisol in an adult population. *Psychoneuroendocrinology* **30**, 857–868 (2005).

101. Vink, J. M., Sadrzadeh, S., Lambalk, C. B. & Boomsma, D. I. Heritability of polycystic ovary syndrome in a Dutch twin-family study. *The Journal of clinical endocrinology and metabolism* **91**, 2100–2104 (2006).

102. Vink, J. M. *et al.* Cervix smear abnormalities: linking pathology data in female twins, their mothers and sisters. *European journal of human genetics: EJHG* **19**, 108–111 (2011).

103. Willemsen, G., van Beijsterveldt, T. C. E. M., van Baal, C. G. C. M., Postma, D. & Boomsma, D. I. Heritability of self-reported asthma and allergy: a study in adult Dutch twins, siblings and parents. *Twin research and human genetics : the official journal of the International Society for Twin Studies* **11**, 132–142 (2008).

104. Van Soelen, I. L. C. *et al.* Heritability of verbal and performance intelligence in a pediatric longitudinal sample. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **14**, 119–128 (2011).

105. Van Leeuwen, M., van den Berg, S. M., Peper, J. S., Hulshoff Pol, H. E. & Boomsma, D. I. Genetic covariance structure of reading, intelligence and memory in children. *Behavior genetics* **39**, 245–254 (2009).

106. De Zeeuw, E. L., van Beijsterveldt, C. E. M., de Geus, E. J. C. & Boomsma, D. I. Twin specific risk factors in primary school achievements. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **15**, 107–115 (2012).

107. Van Bergen, E. *et al.* Why do children read more? The influence of reading ability on voluntary reading practices. *Journal of child psychology and psychiatry, and allied disciplines* **59**, 1205–1214 (2018).

108. Van den Berg, S. M., Posthuma, D. & Boomsma, D. I. A longitudinal genetic study of vocabulary knowledge in adults. *Twin research: the official journal of the International Society for Twin Studies* **7**, 284–291 (2004).

109. Posthuma, D., de Geus, E. J., Bleichrodt, N. & Boomsma, D. I. Twin-singleton differences in intelligence? *Twin research: the official journal of the International Society for Twin Studies* **3**, 83–87 (2000).

110. Van Beijsterveldt, C. E. M., Bartels, M. & Boomsma, D. I. Comparison of naturally conceived and IVF-DZ twins in the Netherlands Twin Registry: a developmental study. *Journal of pregnancy* **2011**, 517614 (2011).

111. Hoekstra, C. *et al.* Body composition, smoking, and spontaneous dizygotic twinning. *Fertility and sterility* **93**, 885–893 (2010).

112. Estourgie-van Burk, G. F., Bartels, M., van Beijsterveldt, T. C. E. M., Delemarre-van de Waal, H. A. & Boomsma, D. I. Body size in five-year-old twins: heritability and comparison to singleton standards. *Twin research and human genetics: the official journal of the International Society for Twin Studies* **9**, 646–655 (2006).

113. Pietiläinen, K. H. *et al.* Distribution and heritability of BMI in Finnish adolescents aged 16y and 17y: a study of 4884 twins and 2509 singletons. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity* **23**, 107–115 (1999).

114. Geels, L. M. *et al.* Increases in alcohol consumption in women and elderly groups: evidence from an epidemiological study. *BMC public health* **13**, 207 (2013).

115. Willemsen, G. et al. The Netherlands Twin Register Biobank: A Resource for Genetic Epidemiological Studies. *Twin Research and Human Genetics* **13**, 231–245 (2010).

116. Costa, P. T. & McCrae, R. R. *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual. Psychological Assessment Resources,* Odessa (Psychological Assessment Resources, 1992).

117. Hoekstra, H., Ormel, J. & de Fruyt, F. *NEO persoonlijkheidsvragenlijsten NEO-PI-R en NEO-FFI. Handleiding.* (Swets & Zeitlinger, 1996).

118. Morey, L. C. *The Personality Assessment Inventory: Professional Manual.* (Psychological Assessment Resources, 1991).

119. Feij, J. A. & van Zuilen, R. W. *Handleiding bij de spanningsbehoeftelijst (SBL).* (Swets & Zeitlinger).

120. Zuckerman, M. Dimensions of sensation seeking. *Journal of Consulting and Clinical Psychology* **36**, 45–52 (1971).

121. Conners, C. K., Erhardt, D. & Sparrow, E. P. *Conners' Adult ADHD Rating Scales (CAARS). Technical Manual.* (Multi-Health Systems Inc., 1999).

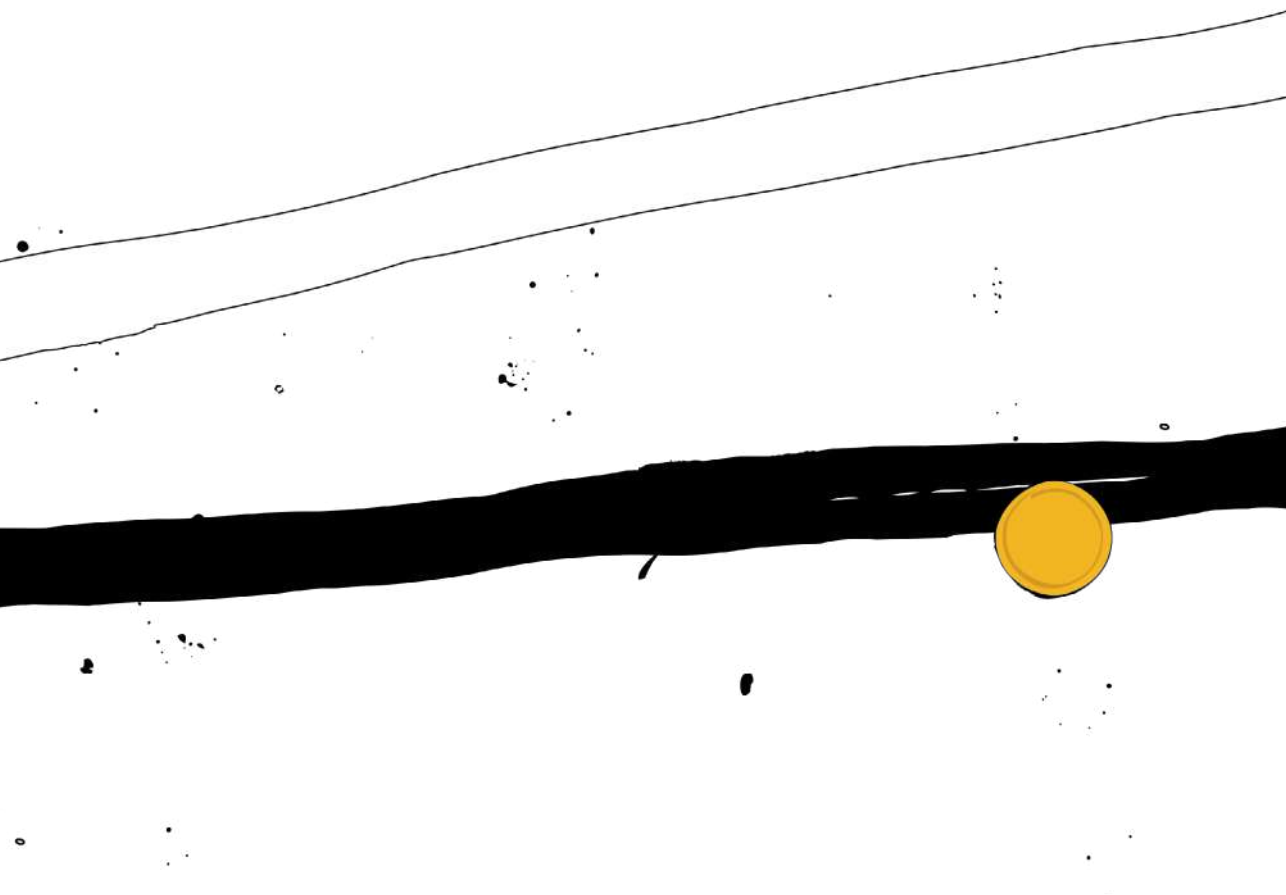
122. Achenbach, T. M. & Rescorla, L. A. *Manual for the ASEBA Adult Forms and Profiles.* (University of Vermont, Research Center for Children, Youth and Families, 2003).

123. Diener, E., Emmons, R. A., Larsen, R. J. & Griffin, S. The Satisfaction With Life Scale. *Journal of personality assessment* **49**, 71–75 (1985).

124. Hughes, M. E., Waite, L. J., Hawkley, L. C. & Cacioppo, J. T. A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. *Research on aging* **26**, 655–672 (2004).

Chapter 11

A twin perspective on what it's like to grow up with or without a co-twin: Personal experiences, scientific, and ethical considerations



This chapter incorporates two book reviews, which I wrote at the invitation of editors of the journal *Twin Research and Human Behavior*. It is exciting to be one of the first to read the books before their publication, and present one's own point of view, and possibly to introduce potential readers to the book and its author. Interestingly, that both books were written by authors who are experienced twin researchers and the twins themselves. The book *Identical Twins: Adult Reflections on the Twinship Experience*[1] was written by Joleen Greenwood in 2018. It is based on an account of the personal experiences of over one hundred identical twins. The book *Deliberately Divided* [2] was written by Nancy Segal in 2021. This book follows the thought-provoking documentaries on twins separated at birth, *The Twinning Reaction and Three Identical Strangers.* For twin researchers, these books are of interest because they provide a perspective on the personal experience of twins, and touch upon important issues of twin research relating to the psychology of twinning, emotions, the involvement of twins in science, ethics and moral dilemmas.

Published as: Odintsova V. (2019). Book review: Identical Twins: Adult Reflections on the Twinship Experience by Joleen Loucks Greenwood (2018). *Twin Research and Human Genetics,* 22(3): 99.

Odintsova, V. (2021). Book review: Deliberately Divided: Inside the Controversial Study of Twins and Triplets Adopted Apart by Nancy Segal (2021). *Twin Research and Human Genetics,* 24(5): 350-351.

"Identical Twins: Adult Reflections on the Twinship Experience" To Feel how it Feels

According to the theory of mind, we have the capacity to feel other people's reality or experiences, and 'to put oneself in another's shoes'. As researchers, we study twins from various points of view, such as their anthropometric measures, cognitive and behavioral characteristics, and health, as well as numerous of other phenotypes. But do we really understand what is it to be an identical twin? What it is like to spend the very first days of life with someone exceptionally similar to oneself in many respects? Or to celebrate every birthday and experience every milestone with said person, while also being compared with him or her? What is this intersubjective dimension of thoughts, feelings and emotions that is extremely unique to twins? It is something very different to just seeing yourself in the mirror. A twin is another person, someone with whom one can talk, share thoughts, endure life's difficulties and just be oneself, as Joleen Loucks Greenwood describes in the book, *Identical Twins: Adult Reflections on the Twinship Experience.*

The twin experience itself is exceptional; the story of even just one twin pair merits its own literary piece. However, the basis of this book is to fully investigate and attempt to understand the personal experiences of 113 identical twins of different ages and stages of life.

In her book, Joleen Greenwood reflects on those questions that are of true concern for twins, but not often taken into consideration by others. For instance, how does this rare occurrence of twinship influence, first of all, the twins' relationships with others — friends, other siblings and partners? Insight into the closeness — and also the strain of a twin pair's relationship that appears when the twinship connection is somehow changed (i.e., by external circumstances or other people's inclusion in the particular dyad) — is provided by many participants in different stories. By analyzing these stories, the author attempts to classify and systemize the twinship experience while also synthesizing its essence. Thus, readers will find many direct citations from interviews with twins throughout the book; however, at the same time, the author groups the answers and defines the twinship experience, sometimes with the help of twin 'jargon' in order to catch the intangible (e.g., 'best friend' with respect to the twin, 'second fiddle' or 'third wheel' with respect to people outside of the twin dyad, 'striving for identity/individuality'). The author also discusses some phenomena, such as the 'buffering effect' from a specific twin position (i.e., the search of additional sources of support in a twinship to aid in coping with a stressful event) and the semantic analysis of words used by twins (e.g., 'never alone', 'always', 'we' instead of 'I', 'unconditional love'). However, Joleen Greenwood notes that the twinship experience can have different periods of togetherness, of sharing relationships with friends, siblings and partners, of distance or separation and even of losing one half of the twin pair. Taking this all into consideration, it is clear that through her book, Joleen captures the dynamic nature of twinship relationships.

The fact that the author is an identical twin herself makes the description more animated and provides a unique look into the intimate world of twins, as it brings in a more personal perspective that makes the text more sincere, and I would also say more vulnerable — you can feel the nerve of the invisible connection between two people brought into the world together.

All of these make reading *Identical Twins* feel as though it is exciting storytelling and convey the experience of what it is like 'never to be alone'. So, although the book is recommended to twin researchers, educators, social workers and psychologists, it is also relevant to a wider readership, including those who are interested in reflections on life through exceptional personal experience. After all, we understand ourselves better through understanding the similarities and differences between others.

"Deliberately Divided". Deluding walls of silence

The book *Deliberately Divided* by Nancy Segal is one of those cases in which the story presented on screen presents a demand for a more detailed analysis. The films, *The Twinning Reaction* and *Three Identical Strangers,* describe the events around a cohort of triplets and twins who were separated at birth in the United States in the 1960s. These events were not only shocking, but raised questions and internal controversy in the audience, including for twin researchers around the world. The five twin pairs and one set of triplets were given up for adoption by their mothers at around 6 months to the adoption agency Louise Wide Services, and separated and given to different families. The adoptive parents were not informed about the existence of co-twins or co-triplets, or the medical history of the biological families, but gave their consent to participate in a study of "child development" conducted by the Child Development Center in New York. The children were followed by researchers who knew all about the circumstances of the separated twins and triplets during the first 12 years of their life. On becoming adults, one by one the children found out the truth and searched for information about their biological parents through DNA testing companies, and via the media. Their emotional first reunion meetings and stories appeared in TV shows and journals. The small scientific group of American researchers known as the Louse Wise Service (LWS) – Child Development Center (CDC) Twin study, led by Professor Peter Neubauer with support of the adoption agency, has created a precedent that may become a stumbling block in collaboration between scientists and research participants around the world and linger for a very long time.

The book can be divided into two parts. The first part includes scrupulously presented material of all the cases of twins and triplets who took part in the LWS-CDC study. The information was collected from all possible sources: interviews, articles from the press and from highly specialized journals read only by small groups of professionals, emails, and telephone conversations. Sometimes participants refused to meet with the author, but still contributed to the story with an explanation for their refusal. We find out details about the biological families, the twins' birth, the first months the twins spent together, the adoptive families and life after separation, the reunion between twins, and life afterwards. This part reads as an engaging detective story. So many parallel lines of discovery are described, many of which are tragedies, of the sensitive reactions to the discovery of being a divided twin, with sometimes a long search for the co-twin, or the loss of one of the twins due to suicide or illness. At times it seems as if an invisible reality inexorably sets in motion certain physiological mechanisms. This determinism is associated with the history of the study participants (often psychopathological diagnoses in biological

parents), vulnerability to emotional and stressful events, and possibly not yet studied long-term consequences of the twin bond and separation.

The second part of the book provides new facts around LWS-CDS and the author's explanations of their conduct based on the understanding of the case in three dimensions – "legalities, framework for research, and moralities".

The genre of the book is different from the documentaries. As Erik Barnouw notes in his book *Documentary: A History of the Non-Fiction Film,* a documentary cannot consider "the truth," but rather the testimony of a fact or event within a social and historical context, when "each selection is an expression of his [documentary filmmaker] point of view"[3]. The film *Three Identical Strangers,* as a documentary, shows that part of reality that is available to the film director with his worldview, the facts he sees, the emotions and questions that have arisen, and all in the context of cinema. The film emerges from the context of already existing television shows, investigative journalism, cherry-picked facts, and triggering.

But is it possible that the scientific community would have given its explanation and reacted to the described events much earlier and before the films were produced? And why didn't this happen? Even if many scientists did not know about the study, there were those who at some stage had tested separated twins and felt uncomfortable *("something that was not quite right", "doesn't feel good"),* or who had heard of the study and discussed it. Why did these representatives of science not initiate an investigation into the events and assumptions of the study, or publish an explanation to the public as soon as the first interviews with triplets appeared in the 1980s, or when Neubauer published *Nature's Thumbprint*[4] in 1990 and explicitly presented the cases of identical twins separated at birth?

In sociology, there is a concept known as "walls of silence", which is a phenomenon when witnesses and participants in events remain silent and do not disclose errors, misconducts, or crimes[5]. It is quite amazing how dense the "walls of silence" can be when dazzling ideas are intertwined, in which many years of scientific research under the rubric of secrecy have been invested, and the emotions of people who discover that their fate has been determined without their awareness are completely disregarded. Nowadays, we are more likely to talk about the ethics of artificial intelligence than about the ethics of human intelligence.

For *Three Identical Strangers,* at least two letters of protest appeared in response to the film and created cracks in the walls of silence. Earlier in her previous book *Twin Mythconceptions*[7]*,* Nancy Segal used a myth-debunking technique regarding twins. And this time the technique was applied with full descriptions of what errors are contained in the use of facts by an authoritative group of scientists in a letter criticizing the film and preventing it from receiving "authoritative" awards. Despite all the resistance, Segal managed to interview or obtain interviews with some of the researchers who participated in the LWS-CDC study, as well as their colleagues and relatives, putting together an incredible archive, the list of which is about 100 pages at the end of the book. When the question "How could this happen?" arises, an unbiased position of an expert who takes the courage to answer the question and is not afraid to be in the cross-hairs, is extremely important.

Nancy Segal is an extraordinary expert who has been devoted to twin studies for most of her professional life. Thanks to this experience, her presentation encompasses much more than a single research project. She knows the detailed stories of hundreds of twins from the unique cohorts of twins who were reared apart or switched-at-birth[6]. Her talent as a writer has presented readers with several books on these topics, and perhaps no one could better tell the story of the LWS-CDC study, which has shocked the viewers of the TV shows and films about the study. It is quite easy to take the side of the creator of the hottest news. But truth requires sensitivity, expertise in the field of the research object, and Socrates' ability to catch and explain contradictions and change optics, to reveal the motivation of all participants in the described events. In this book, Segal managed to accomplish those feats.

What distinguishes this book from Segal's previous ones is precisely the analytical search for the motivations of not only the twins, but also the researchers themselves and the adoption agency involved. This is where we learn much more about the other side of twin research – often invisible, and impersonally presented in mass media as *"scientists have made a discovery..."* The LWS-CDC study was not initiated with a scientific perspective, but from the belief that to grow up with a twin (being also from a family with psychiatric anamnesis) could be harmful to children. When Segal seeks the researchers' motivation, she provides an excellent sociological description of the scientific context of the time, and approaches it from the mindset of those surrounded by ideas circulating about psychoanalysis, nature and nurture, genetics and the environment. This was a time just after the discoveries of molecular biology that irrevocably turned science towards zooming in on the chemistry of human life. In fact, this book is unique in its analysis of human delusions – not just philistine delusions due to ignorance or emotional upheavals, but highly intellectual, argued delusions regarding the Machiavellian postulate "the end justifies the means". This means that it finds a place on the shelf not only of scientific literature, but also of philosophical and legal literature.

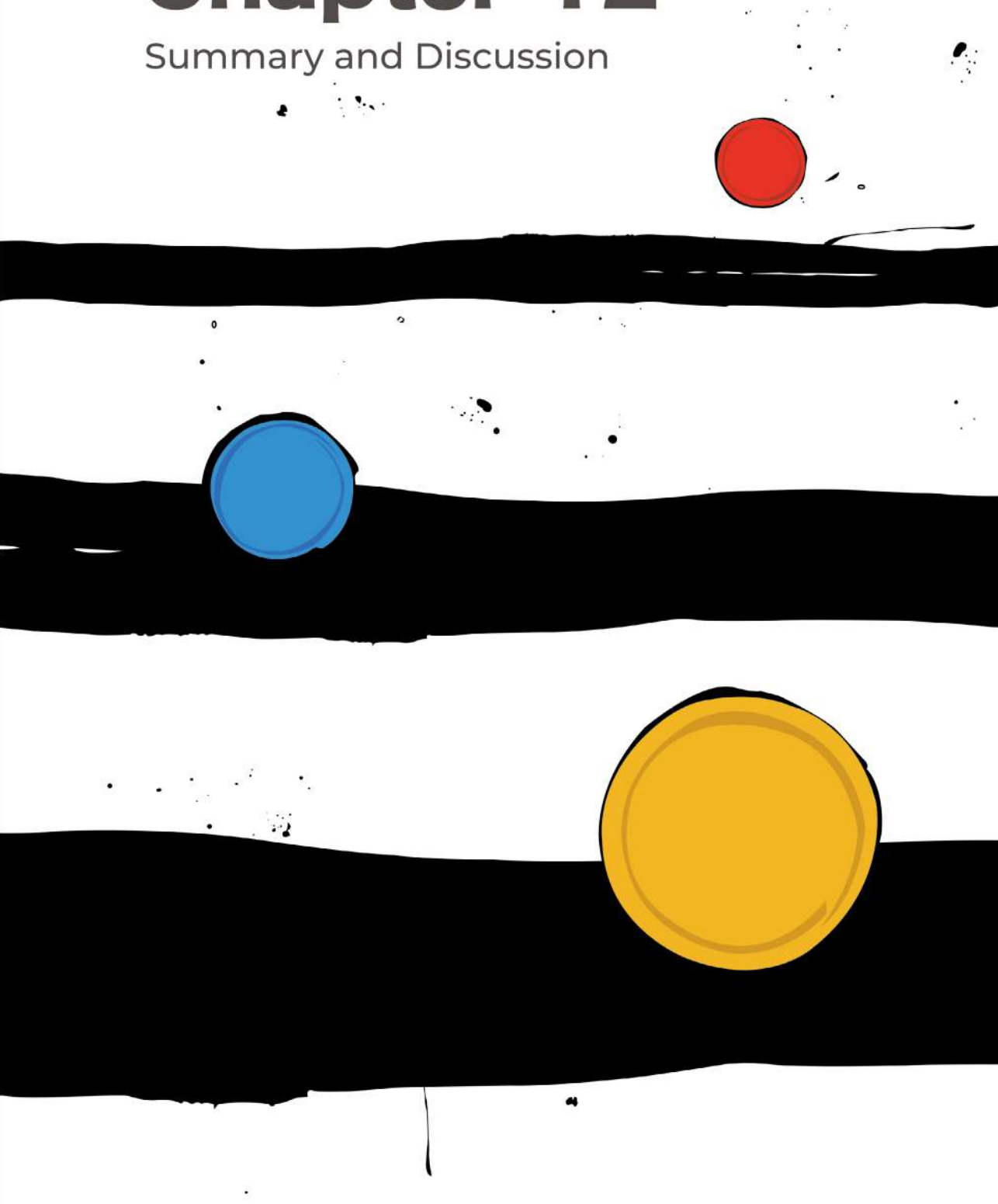
The scientific value of twins is undeniable. The classical twin study is a very "clever and elegant" study design, based on two types of twins – identical and fraternal (mono- and dizygotic), developing in parallel from the moment of conception. Twins around the world are active in scientific research and have contributed to many discoveries in the fields of psychopathology, genetics, behavior and development. As Segal notes, "Twins usually enjoy being in research because they are eager to learn more about themselves and are happy to advance scientific understanding in the process"[2].

References

1. Greenwood, J.L. Identical Twins: Adult Reflections on the Twinship Experience. (Lexington Books, 2018).
2. Segal, N.L. Deliberately Divided: inside the controversial study of twins and triplets adopted apart. (Rowman & Littlefield Publishing group, 2021).
3. Barnouw E. Documentary: A History of the Non-fiction Film (New York: Oxford University Press, 1993), p. 287.
4. Neubauer P, Neubauer A. Nature's Thumbprint. The New Genetics of Personality. (New York: Columbia University Press, 1990).
5. Cohen S. States of denial: Knowing about atrocities and suffering. (Cambridge: Polity, 2001).
6. Segal, N.L. Born Together – Reared Apart: The Landmark Minnesota Twin Study. (Cambridge, MA: Harvard University Press, 2012).
7. Segal, N. L. Twin Mythconceptions: False Beliefs, Fables, and Facts about Twins. (New York: Elsevier, 2017).

Chapter 12

Summary and Discussion



The research presented in this thesis broadly covers the genetics and epigenetics of early life development. It describes several traits and exposures in association with early life characteristics and DNA methylation, the prediction of traits and exposures based on genetic and epigenetic data. Furthermore, this thesis describes the methodology associated with twin data in population-based and biomarker studies. I studied the early life characteristics and DNA methylation in a large prospective population-based cohort of the Netherlands Twin Register, with, in some chapters, additional data from the prospective population-based Avon Longitudinal Study of Parents and Children (ALSPAC) in the United Kingdom.

The thesis is divided in three parts. The **first part** is dedicated to early life traits and exposures.

Chapter 2 describes an etiological study of variation in one-minute and five-minute Apgar scores. Apgar stands for Appearance, Pulse, Grimace, Activity, and Respiration. The score is used to rapidly assess the status of newborns after birth, and to determine if respiratory support is needed. Variance decomposition in a sample of 5,181 twin pairs identified significant estimates of shared and non-shared environmental variance, but no influence of genetic factors. I found that gestational age and two twin-specific characteristics, namely zygosity and birth order, have an effect on Apgar scores. Other characteristics have different effects in the first and second born twins. Low birthweight, cephalic fetal presentation at birth, and mode of delivery with caesarean section, forceps or vacuum extraction increase the probability of having low Apgar scores in first-born twin. A larger time period between delivery of twins, non-cephalic fetal presentation at birth and vaginal mode of delivery increase the probability of having low Apgar scores in second-born twins. Our findings did not confirm the effects of several characteristics previously associated with Apgar scores, such as sex of the child, mode of conception and parental characteristics at birth (maternal and paternal age, mother's BMI).

In **Chapter 3** I introduced the concept of DNA methylation, one of the molecular mechanisms of gene expression regulation. In a review of the literature, I discussed its role in early development and how it manifests in the effects of prenatal maternal smoking exposure. Over 6 thousand CpG sites across the genome showed significant changes in DNA methylation level in children exposed to prenatal maternal smoking, with the strongest associations with CpG sites in the *AHRR gene*[1]*.* These changes show a long-lasting character, although with increasing age the effects decrease. The relevant genes are predominantly involved in developmental processes, such as growth and anatomical development. Several studies detected an association between maternal smoking, epigenome, and such outcomes as obesity[2], inadequate immune response[3], and neurodevelopmental disorders[4].

Chapter 4 reports the findings from my EWAS of breastfeeding in 1,006 children of whom 517 children were 5-9 years old (mean age 7.9, sd 1.14), and 489 children were 10-12 years old (mean age 11.2, sd 0.73). In children younger than 10 years, I found that buccal cell DNA methylation at CpG sites in *ZNF232, MUCL1, DSCR3,* and *ATG10* genes and at three intergenic positions in the genome was associated with initiation

of breastfeeding. In children older than 10 years methylation differences at these CpGs were smaller and non-significant. These differentially methylated positions were associated with mQTLs and are known to be associated with prenatal maternal smoking and infertility. The genes related to these CpGs were associated in GWASs with family history of Alzheimer's disease, bone mineral density, metabolic disorders, regulation of inflammation (cytokines), and cell signaling processes. The findings did not replicate in an independent NTR sample of 98 children of the same age, with buccal DNA methylation data. No nearby sites were associated with breastfeeding in the ALSPAC sample of 938 children of the same age with DNA methylation in blood.

Chapter 5 summarizes the associations of early life characteristics with handedness, defined as left-handedness, righthandedness and non-righthandedness (that combines left- and mixed-handedness). Twin-specific characteristics, such as zygosity, chorionicity, amnionicity, birth order, and time between birth of first and second twin, did not play a role in handedness in my study of 37,495 NTR twins. Being male and preterm increased the probability of being left-handed or mixed-handed. Effects of other characteristics, such as mother's and father's handedness, being breastfed, differed for left-handedness and mixed-handedness. Parental handedness and being breastfed is associated with left-handedness. Other early life characteristics, i.e., year of birth, mother's and father's age, mode of conception, maternal smoking and maternal stress during pregnancy, being born in summer, fetal presentation, mode of delivery, birth weight, and Apgar score were not associated with handedness. Mixed-handedness was associated with adverse future outcomes in neurodevelopmental and externalizing behaviors.

Following a large genome-wide association study of handedness[5], we were the first to perform an epigenome-wide association study of left-handedness combining two large datasets from the NTR and ALSPAC adults (*N*=3,914) with DNA methylation in blood, in **Chapter 6**. We found two differentially methylated regions in *BLCAP* and *NNAT* on chromosome 20 and in *IAH1* on chromosome 2 associated with left-handedness. The genes located in the region on chromosome 20 are imprinted in the brain and are involved in brain development and nervous system functions. The hypomethylation of the region in blood cells in left-handed individuals suggests that genomic imprinting may play a role in handedness. The CpGs from this region have previously been associated with myalgic encephalomyelitis, chronic fatigue syndrome, preterm birth, obesity, metabolic parameters, and arm fat mass (DXA scan measurement). The region on chromosome 2 contains the protein-coding gene that is associated with several diseases, such as neonatal inflammatory skin disease, bowel disease, and a disease due to an inborn error of leucine metabolism, and CpGs that were previously associated with psychiatric diseases including schizophrenia. Overall, the observed effects of left-handedness on DNA methylation were small. Therefore, we could not conclude that DNA methylation is associated with left-handedness in blood cells or buccal cells.

In the second part of my thesis I started in **Chapter 7** with a review of the candidate gene, genome-wide association, epigenome-wide association, and other omics domain studies of aggression and related phenotypes. For this literature review, I used an automated literature selection tool. Our analysis of manual and automated procedures revealed that the automated literature selection tool resulted in a decrease of ~40% in

workload. In addition, the automated tool detected literature not found by the traditional approaches for literature searches.

We found that only 4 studies out of 18 genome-wide association studies of aggression reported genome-wide significant hits. We identified five EWASs, which detected a number of differentially methylated loci in blood and buccal cell DNA related to aggression. The largest EWAS meta-analysis reported 13 genome-wide significant CpGs associated with aggression[6]. However, the DNA methylation score explained only 0.29% of the variance in aggression in adults[6]. Polygenic scores, in comparison, explained between 0.04% and 0.44% of the variance in aggression in 7-year old children[7].

The prediction based on polygenic scores (PGS) and by DNA methylation scores was analyzed more in detail in **Chapter 8**. Blood DNA methylation scores were constructed based on summary statistics from the largest EWASs for several traits. The DNA methylation scores predicted these traits independently from PGSs. In an adult NTR sample (N=2,431), DNA methylation scores outperformed PGSs for a number of traits, including smoking status (R2 54.4% vs 0.16%), former smoking (R2 15.5% vs 0.5%), prenatal maternal smoking (R2 16.9% vs 0.5%), and BMI (R2 6.4% vs 5.5.%), but not birth weight (R2 0.4% vs 1.5%). These same traits (with the exception of own smoking status) were also analyzed in buccal cell DNA in 9-year old NTR children (N=1,128), but were not predictive of any of these outcomes.

In the **third part** of the thesis, I discussed methodological tools and developments of new resources in the field of molecular genetics studies using twin data.

In **Chapter 9**, I described the use of twin resources for guiding post-GWAS studies, and the potential for integration of twin registers with biobanks and other resources to boost research, rising to new challenges in the field of gene finding, pharmacogenomics, and search of disease biomarkers for medicine. I discussed the opportunities to enhance phenotyping through record linkage from multiple sources and the addition of more biological data collections to twin register biobanks. Several critical issues that can influence the twin research findings were discussed. Zygosity misclassification can play a crucial role in the heritability estimates within a twin study reducing the overall heritability estimate. The processing of batches of biological samples together with family clustering calls for statistical correction procedures, as shown in this chapter. At the end of **Chapter 9**, I presented several examples of successful consortia collaborations, which by pooling data from a large number of twin cohorts may advance the analysis of the genetic and environmental underpinnings of human complex phenotypes.

In **Chapter 10**, the comparison of 10,176 multiples with 2,142 siblings demonstrated that twins do not differ from their singleton siblings with respect to a wide range of behavioral and psychological traits, biomarkers, and diseases. One observed difference concerned body mass index and height. Twins remained smaller and lighter compared to their singleton siblings, even as adults, with a BMI 1 point lower than singletons. In the literature congenital disorders are identified as the most important source of twin-singleton differences, with a twin or multiple pregnancy at an increased risk of prenatal mortality and congenital abnormalities. Quotes from the twins on unique twinning experiences and their perception of the multiple status as something special are included at the end of **Chapter 10**.

Observations on the twins' personal experiences are the focus of **11 Chapter**. It describes the dynamic nature of twinship relationships through the lifespan and discusses the potential harmful effects of separation of twins that happened in the past, and the ethics of inclusion of twins in research.

General discussion

In this section, I would like to discuss some findings from different chapters of my thesis, sometimes connecting the dots on common issues from different studies. I will reflect on questions related to association versus causality, the importance of replication of discovery study findings, and polygenic and DNA methylation prediction of complex traits and exposures. I will end with some reflections on future directions in the field of complex trait genetics and personalized epigenomics.

Early life factors and epigenome

In this section, I discuss my findings on associations between early life characteristics, and epigenome-wide association studies.

Studies into the relationship of adverse early life environment and later health outcomes in life, especially noncommunicable disease, are not novel[8]. Many prenatal and early postnatal factors are related to adverse outcomes in epidemiological studies. These factors include prenatal maternal smoking[9], heavy alcohol consumption[10], prenatal undernutrition[11], folic acid insufficiency[12], low socioeconomic status[13], birth complications[14,15], maternal birth stress during pregnancy[16], and absence of breastfeeding[17].

In my thesis, I studied a set of prenatal, perinatal, and postnatal characteristics in association with Apgar scores and non-righthandedness. We found that some characteristics, such as shorter gestational age, lower birth weight, breach or horizontal fetal presentation, delivery tactics in complicated deliveries, monozygosity, being a second-born and longer delivery time were related to poor newborn status at birth, as reflected in Apgar scores. Other characteristics, including mother's and father's age, mother's BMI, mode of conception, and maternal smoking during pregnancy, were not associated with Apgar scores. Both left-handedness and mixed-handedness were associated with shorter gestational age and male sex.

Associations between early life characteristics and non-righthandedness are not necessarily causal. The association with sex and parental handedness indicates that a genetic component plays a role in non-righthandedness. The association of handedness with gestational age and left-handedness with breastfeeding could reflect an influence of some other (latent) factor(s). Similarly, the relationship between mixed-handedness and adverse neurodevelopment outcomes and externalizing problems does not necessarily imply a causal link between those outcomes and mixed-handedness.

Our findings on relationships between early life factors and outcomes are in line with the so-called Barker hypothesis that relates prenatal adverse factors that affect the pregnant woman, and consequently her fetus, with several diseases in adulthood[18–20], and with the concept of developmental plasticity. Developmental plasticity relates the adverse prenatal environment with the metabolic processes in fetus and susceptibility to disease[21].

The epigenetic mechanisms in this context lead to long-term changes in gene expression through chemical modifications to, or alterations of, the packaging of DNA, such that the capacity for transcriptional regulation is altered[22]. The DNA sequence is static, although its expression is subject to change. Epigenetics can reflect a variety of factors including the DNA sequence, an exposure (e.g., intrauterine environment), an inherited imprinting, a consequence of a disease, disease-associated factors, or random effects. Some studies have provided evidence that DNA methylation mediates effects of adverse early life factors, such as gestational diabetes[23] and maternal smoking during pregnancy[24], on poor outcomes later in life. In this thesis, I reviewed epigenome-wide association studies of maternal smoking during pregnancy and performed epigenome-wide association discovery studies of breastfeeding and left-handedness. These EWASs were exploratory and focused primarily on establishing associations.

In my review on maternal smoking during pregnancy, I described the comprehensive meta-analysis of Joubert et al.[1], which detected over 6 thousand differentially methylated CpGs across the genome in children whose mothers smoked during pregnancy. Following the meta-analysis of Joubert et al. (2016), Sikdar et al.[25] addressed the question of whether the differentially methylated CpGs associated with prenatal maternal smoking in newborns overlap with differentially methylated CpGs associated with own smoking in adults. A substantial number (1709) of differentially methylated CpGs overlapped, and a large number of CpGs (3838) were unique for prenatal maternal smoking. Wiklund et al. studied the potential mediatory role of DNA methylation in the association of maternal smoking during pregnancy and adverse offspring health outcomes[26]. They addressed the issue of whether methylation markers have causal effects on disease outcomes in offspring between the ages of 16 and 48 years with Mendelian randomization and mediation analysis. In this case, Mendelian randomization was facilitated by instrumental genetic variants that are known to be strongly associated with DNA methylation, known as methylation quantitative trait loci (mQTLs)[27]. Their analysis provided evidence for a causal role of four maternal smoking-related CpGs on an increased risk of inflammatory bowel disease or schizophrenia.

In children of around eight years old, I reported seven CpGs associated with breastfeeding assessed in buccal cell DNA. The association between DNA methylation and being breastfed may be due to a causal effect of breastfeeding on DNA methylation or confounding effects[28]. Confounding prenatal and postnatal factors may include low sucking ability of premature, low birthweight newborns, mother's decision not to initiate breastfeeding due to complications at delivery, or the difficulty of having two infants to breastfeed. As was shown by Merjonen et al.[29] in an analysis of NTR data, 70% of individual differences in initiation of breastfeeding are explained by genetic influences. Thus, DNA methylation differences could reflect the genetic component of the breastfeeding initiation transmitted from mothers to children. Walker-Short et al. tested methylation in cord blood and peripheral blood in children and found three CpGs significantly associated with breastfeeding at birth, infancy and childhood[30]. According to Walker-Short et al., the explanation may be that the child inherited methylation marks associated with the breastfeeding exposure, and that this exposure is associated with the mother's decision to breastfeed. Subsequent studies on epigenetics of breastfeeding effects should include more detailed data on mother's genetics and health status,

her decision to initiate breastfeeding, data on breast milk compounds, maybe though metabolomics studies, and data on infant diet.

The low estimate of the heritability of handedness (~25%[31]) suggests that the etiology of handedness involves mechanisms beyond DNA sequence. Our joint meta-analysis of NTR and ALSPAC data detected two differentially methylated regions associated with left-handedness. The CpGs from this EWAS meta-analysis that are located close to the left-handedness-related SNPs from GWAS meta-analysis[5] have stronger associations with left-handedness and are not driven by DNA methylation quantitative trait loci (mQTLs). This indicates that both DNA methylation and genetic variants at these loci are associated with left-handedness. However, the observed effects of CpGs were small. Future epigenetic studies in other tissues related to the nervous system may shed light on the etiology of handedness. Given that laterality and handedness appear very early in prenatal development[32,33], we expect that some biomarkers of laterality in peripheral tissues will be found with larger samples, larger coverage of the epigenome, and measurements of DNA methylation at early stages of fetal development.

The relationship of left-handedness and breastfeeding that we found has also been reported in several other studies[34,35]. As we performed the EWASs of both breastfeeding and left-handedness, we specifically looked at DNA methylation that was associated to both phenotypes. If we look at DNA methylation in buccal cells in association with breastfeeding and left-handedness in NTR children of around 9 years old, we find one common CpG in top 500 CpGs (cg17337463, chr1, *RAP1A;* β_{LH} = 0.027, *p*-value 1.9×10^{-05}; $\beta_{breastfeeding}$ = -0.026, *p*-value=0.0005).

Considering DNA methylation in different tissues – buccal cells in NTR children in association with breastfeeding, and peripheral blood samples in adults in the left-handedness meta-analysis – we found no common CpGs in the top 1000 CpG's summary statistics from the EWASs. Thus, we did not find evidence for common DNA methylation differences associated with breastfeeding and left-handedness, in peripheral tissues. Further studies are needed to explore the genetic and DNA methylation architecture behind the relationship between breastfeeding and handedness.

Replication and reliability

In this section, I discuss some issues of replication of epigenome-wide association studies results. An EWAS is generally more complex than a GWAS, because the epigenome is dynamic and differs across tissues and over time. Thus, there are more factors influencing EWAS results than there are influencing GWAS results. This may be one reason that different epigenetic studies do not always replicate. Van Rooij et al. discussed the EWASs of current smoking, which differed with respect to the results regarding number of significant CpGs across EWAS studies (taking into account sample size)[36]. Replication in epigenetic investigations is as important as it is in genetic studies. However, diversity in the EWASs' methodologies (DNA source, age at biosample collections, batch effects, correction for confounders etc.) complicate the generation of reproducible EWAS results. As I discussed in my paper on establishing a twin register, biomarker assessments may include batch effects, small sample size per batch, and dependent observations in the batches (i.e., family members). It is, however, also important to devote attention and

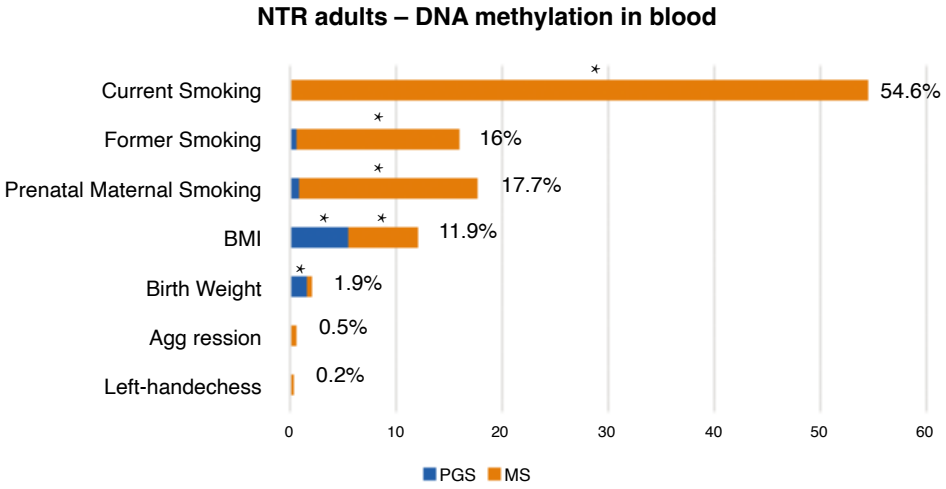
resources to carefully assessing age at sample collection, to collecting multiple tissues, e.g., buccal cells in addition to blood, and to the measurement of cell composition, as was done in NTR biobank[37].

The differences in the adjustment for covariates and source of DNA may explain differences in the replication of EWASs of breastfeeding. Walker-Short et al.[30] did not replicate our findings in buccal cell DNA methylation in the sample of children with DNA methylation in peripheral blood. They found similar associations with CpGs located nearby or in *LEP,* detected in candidate gene study analyzing DNA from buccal cells[38], and in an EWAS analyzing DNA isolated from peripheral blood[39]. Our study failed to replicate previously reported CpGs located nearby *LEP*[39,40]*, IL4R*[41]*,* and *CDRN2A*[42], but replicated four other hits that have been detected in an EWAS using peripheral blood[43].

As shown by Naumova et al. in an EWAS of breastfeeding, about 10% of the variability in methylation level of CpGs is associated with the variability in cell-type composition in peripheral blood[43]. Most EWASs included cell counts, technical variables, sex, and age at DNA collection, and also other covariates that are strongly correlated with the phenotype of interest, or that are known to affect DNA methylation. Our analysis of breastfeeding included a wide set of covariates such as maternal smoking during pregnancy, gestational age, SES, maternal pre-pregnancy BMI, and maternal age at delivery. However, it is unlikely that a wide set of covariates is always available. In general, covariates included in EWASs of breastfeeding differ among studies[39,42]. Another issue in replication of our findings concerns the type of tissue for DNA methylation. The associations between DNA methylation in buccal cells and breastfeeding, which we found, were not replicated in the ALSPAC cohort with DNA methylation in peripheral blood.

We expected that left-handedness would be associated with similar DNA methylation differences across multiple tissues given the potentially early processes of lateralization in early development[32]. We attempted to replicate our meta-analysis results in both peripheral blood and buccal cell DNA methylation samples in ALSPAC and NTR with EWASs performed with the same set of covariates. However, replication was not seen, which may be attributable to a lack of power and (or) the difference in tissue. The difficulties in applying cross-tissue DNA methylation analysis were also demonstrated by the lack of predictive value of DNA methylation scores, generated with different DNA source in discovery and target studies. Currently, there is no clear evidence that DNA methylation changes respond to environmental exposures in a similar way across tissues[44].

Another important issue regarding the reproducibility of findings is statistical power, which is related to sample size and small effects of single CpGs. Power calculations for EPIC array studies revealed that >80% power to detect an effect of 2% requires a sample size of 500 cases and 500 controls for complex phenotypes[45]. Differentially methylated position analysis (site-by-site) leads to underpowered analysis, given the small effects at a single CpG. To increase sample size, meta-analysis is commonly applied. This involves aggregating results from different populations with measurement of DNA methylation obtained at different time-points, and provides more power and generalizability of results. To solve the problem of small effect sizes in EWASs, different techniques are applied, such as differentially methylated regions (DMR) analysis and the use of DNA methylation scores. The analysis of differentially methylated regions



Trait	*N* CpGs in MS	*N* sample in target sample (NTR)	*N* in discovery study	Tissue in discovery study
Smoking	24	1,914	9,389	blood
Prenatal Maternal Smoking	607	720	4,994	cord blood
BMI	412	2,410	5,387	blood
Birth weight	934	2,040	8,825	cord blood
Aggression	745	2,056	12,375	blood
Left-handedness	14	2,198	1,058	blood

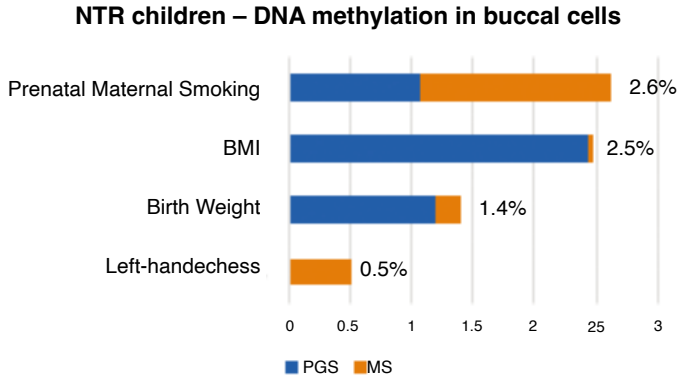
Figure 1. Multi-omics prediction by polygenic scores and DNA methylation scores in the Netherlands Twin Register adult dataset with DNA methylation data in blood (*N* ranged from 720 to 2,410)

* indicates significant prediction of the trait;
PGS, polygenic score; MS, DNA methylation score.

combines statistics from different CpGs located within a certain region[46,47]. The construction of DNA methylation scores based on summary statistics from earlier studies and testing its predictive ability is another approach of combining the measurements across the genome, and increase power[48].

We applied meta-analysis and DMR analysis in the study of left-handedness, adjusting for dependencies between the sites and uncertainty in the EWAS effects. The detected left-handedness-associated regions are results for further study of laterality and handedness in animal and human models. DNA methylation scores were predictive for several traits, but not for left-handedness, as discussed in the following section.

Summarizing, the epigenetic analyses of breastfeeding and left-handedness illustrate the common problem of obtaining reproducible EWAS results[44]. Furthermore, the field of EWASs faces challenges regarding standardization of analytical methods and power. The questions of how to reduce the heterogeneity and which covariates to include in



Trait	*N* CpGs in MS	*N* sample in target sample (NTR)	*N* in discovery study	Tissue in discovery study
Prenatal Maternal Smoking	606	547	4,994	cord blood
BMI	183	1,072	5,387	blood
Birth weight	958	1,070	8,825	cord blood
Left-handedness	12	799	742	blood

Figure 2. Multi-omics prediction by polygenic scores and DNA methylation scores in the Netherlands Twin Register child dataset with DNA methylation data in buccal cells (*N* ranged from 547 to 1,072)

PGS, polygenic score; MS, DNA methylation score.

the EWAS analyses remain open. Correcting for such traits as BMI and smoking substantially changes the results, and is often recommended in analysis of complex traits and exposures in order to remove their variation from the data[36]. Studies that report results at a lenient thresholds for the probability of making a type I error when analyzing small sample size should be interpreted with caution[45].

Prediction by polygenic scores and DNA methylation scores

The aim of this section is to discuss the findings on polygenic scores (PGS) and DNA methylation scores (MS), as described in different chapters in this thesis: prenatal maternal smoking, birth weight, body mass index (BMI), current and former smoking, left-handedness, and aggression with reports from the studies in the NTR[6,7].

Prediction of an individual's phenotype or exposure from genetic and epigenetic data is one of the great expectations in genomics and precision medicine[49,50]. The predictive value of a genetic or epigenetic predictor may be interpreted as the extent to which the predictor predicts an individual outcome. The proportion of explained variance is used to evaluate the performance of the predictor at population level. By combining the effects of multiple SNPs or CpGs into a single score, a larger proportion of variance is explained compared to the variance that is captured by a single SNP or a single CpG. PGSs and

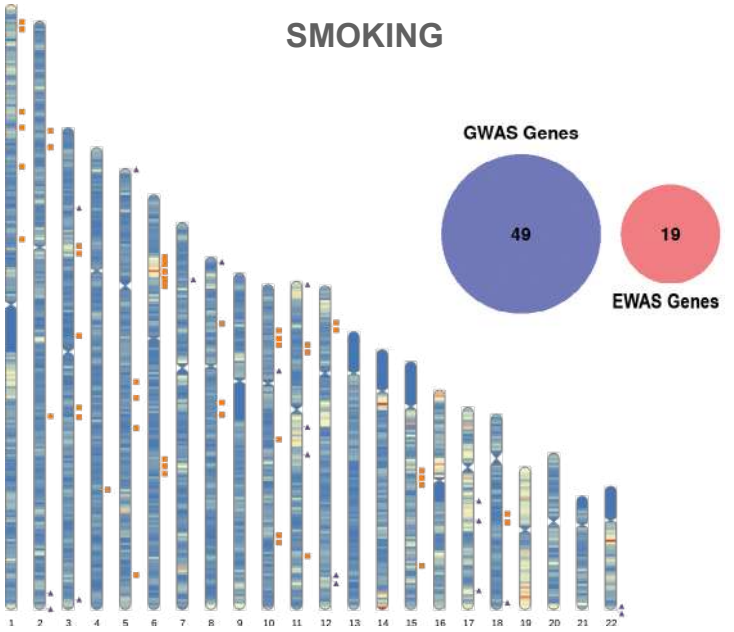


Figure 3. Mapping, overlap of genes identified by GWAS and EWAS of smoking

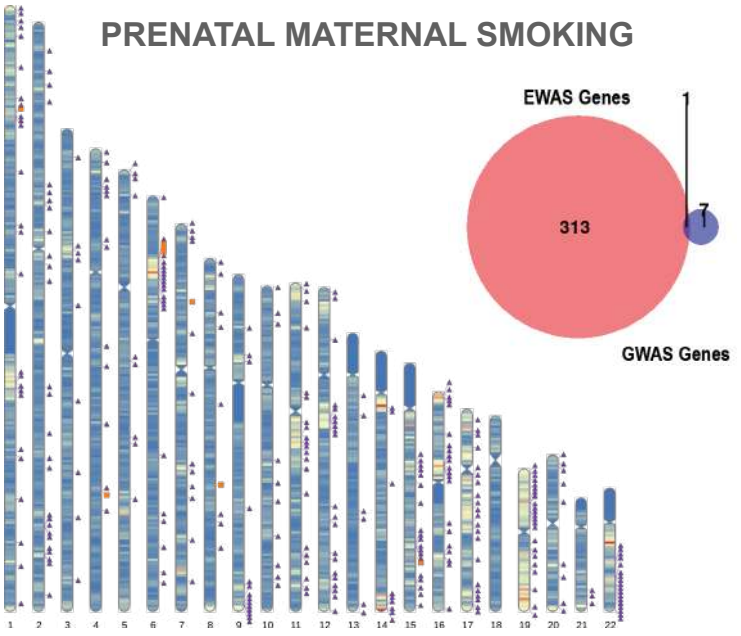


Figure 4. Mapping, overlap of genes identified by GWAS and EWAS of prenatal maternal smoking

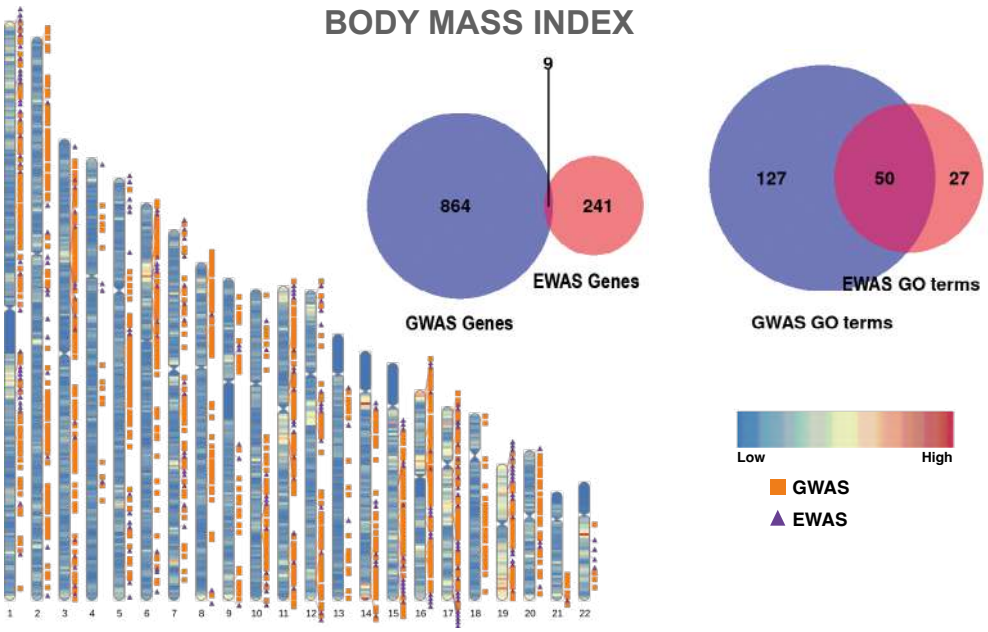


Figure 5. Mapping, overlap of genes identified by GWAS and EWAS of body mass index and overlaps in Gene Ontology (GO) terms "GO Biological Process"

Note in Figure 3-5 that the chromosomes are colored by gene density where blue indicates low density, white intermediate density and red high density.

MSs translate GWAS and EWAS results back to the individual level, by summing the association effects of all loci or all CpG's, that are observed in an individual. PGSs measure genetic susceptibility for a certain outcome at the individual level. MSs extend beyond genetic susceptibility, and may capture environmental and stochastic influences, and the effect of the trait and exposure on the DNA methylation.

Both PGSs and MSs can contribute independently to the prediction of a trait. Both depend on the effect sizes of the associations in the discovery study, which in turn depend on sample size and other factors (e.g., ethnicity, sex, reference set, LD structure, etc.). Currently PGSs are sufficiently predictive for human height[51,52] for example, and MS predictions work reasonably well for age[53] and smoking[54,55]. Previous studies also have shown that for BMI and mortality, DNA methylation scores based on multiple CpGs from EWASs show better predictive ability than currently available polygenic scores[56,57].

We combined genomic and methylation data for several traits and exposures in two datasets of adults with blood DNA methylation data and children with buccal cell DNA methylation data in the Netherlands Twin Register. The syntheses of the polygenic and DNA methylation predictions are presented in **Figures 1** and **2**.

In addition, to add to the discussion of these findings, we (Odintsova and Pool) did a follow-up of our studies on PGS and MS prediction, and performed Gene Ontology (GO)[58]

enrichment for traits and exposures, of which the PGS and MS explained a statistically significant proportion of variance (i.e., own smoking, prenatal maternal smoking, and body mass index). Note that the results of this follow-up are preliminary at this stage. Our main question was: what biological relevance may polygenic and methylation scores contain, and do they cover the same part of the genome or do they differ in mapping the genes? We performed the analysis on SNPs from the GWAS by thresholding and clumping, and CpGs from EWAS studies by thresholding and pruning. The genome-wide significant (p-value < 5×10^{-8}) index SNPs for each locus resulting from the clumping of the GWAS summary statistics for each phenotype were annotated using the Variant Effect Predictor of Ensembl[59,60]. To generate a complete set of SNP annotations, the index SNPs were also queried for the nearest gene using Ensembl BioMart[61] and the resulting annotations were added to the VEP outcomes. The CpGs from the EWAS summary statistics (for smoking at $p < 1\times10^{-7}$, pruned, for prenatal maternal smoking at $p < 1\times10^{-7}$, unpruned, and for BMI at $p < 1\times10^{-1}$, unpruned) were annotated using the Infinium annotation file.

The GWAS gene sets were *(1)* compared with respect to composition to the above EWAS gene sets. The gene sets for each EWAS-phenotype and GWAS-phenotype combination were subsequently taken as input for Gene Ontology (GO) term enrichment analysis for the GO "Biological Process" category[62–64]. For each phenotype the resulting sets of significantly enriched (FDR p-value < 0.05) GWAS and EWAS GO terms were *(2)* again compared with respect to composition.

The results are presented in **Figures 3-5**, which give the genetic locations of the SNPs and CpGs annotated to the SNPs and CpGs for own smoking, prenatal maternal smoking, and BMI; the overlaps of gene sets for all phenotypes, and the overlap of GO terms for BMI. **Supplemental Table 1** shows the overlapping genes, and **Supplemental Table 2** shows the overlapping GO terms for BMI.

The methylation scores were generated based on weights from the discovery study in cord blood, and blood. In the target sample of NTR adults methylation scores (blood) outperformed PGS for own smoking, prenatal maternal smoking, and BMI (**Figure 1**). The methylation scores did not predict birth weight, but PGS did.

Most variance in the trait/exposure is explained by MS for own smoking and prenatal maternal smoking. Tobacco exposure is considered an environmental exposure that modifies DNA methylation, and multiple DNA methylation signatures were identified, and scores were created for smoking[55,65–69] and prenatal maternal smoking[1,67,70]. Gao et al. demonstrated that while mQTLs for smoking-related probes have been detected, a substantial number of CpGs are independent of nearby SNPs[71]. Our follow-up analysis showed that top GWAS and EWAS gene sets had no common genes, and no common biological processes GO terms for smoking. One common gene was detected for prenatal maternal smoking exposure: a protein-coding gene *MAML3* (mastermind like transcriptional coactivator 3) on chromosome 4. We were the first to create the PGS for maternal smoking based on the weights from the UK Biobank. Interestingly, the PGS added 0.8% to the explained variance in the trait.

The MS derived based on DNA isolated form blood in adults explained a significant proportion of variance in BMI and outperformed the PGS. Together the PRS and the MS

accounted for 12% of the BMI variance. Both scores contributed independently in our prediction analysis. Nine common genes were found in the GWAS and EWAS gene sets in our follow-up analysis (**Supplemental Table 1**), and fifty common GO biological process terms (**Supplemental Table 2**). Around 30% of GO terms were related to metabolism.

The variance in birth weight was explained by PGS, but not by MS. However, there is evidence that differentially methylated patterns in adults are related to birth weight[72] and preterm birth[73] (resulting in low birth weight). We expect that the prediction will improve with the publication of discovery EWASs of birth weight in adult populations, and with the construction of MSs using the same source of DNA for methylation profiles in discovery and target samples. BMI and birth weight are physical characteristics, with heritabilities between 47-90% for BMI and 26% for birth weight. BMI also has a behavioral component, that is confirmed by enrichment with genes related to brain and the central nervous system[74,75]. In addition, birth weight-associated CpGs have been related to prenatal maternal smoking and mother's BMI before pregnancy[76].

The prediction of behavioral traits was not significant in our studies. There were no other reports on MSs of left-handedness and aggression in the literature, in addition to our NTR studies. PGSs were generated using the largest meta-analysis to date as discovery studies, including around 1.5 million individuals for handedness5, and 87.5 thousand individuals for aggression[77]. However, the variance in these traits explained by PGS and MS separately and together did not reach 1%. Recently, Ocklenburg et al.[78] constructed PGS of handedness based on the GWAS by de Kovel et al.[79] in 296 adults. PGS was significantly associated with left-handedness assessed by a lateralization quotient that combined hand preferences for various activities (incremental R^2=2.6%). PGSs generally do not explain much variance of psychiatric diseases and related complex traits, due to factors such as heterogeneity of the phenotype, and power issues[80,81].

According to Wittgenstein's Rule-following Paradox, where he states: "no course of action could be determined by a rule, because every course of action can be made out to accord with the rule"[82], it is hard to define conclusively how much variance should be explained to establish an accurate predictor in the context of PGS and MS prediction. For MSs generated in our studies, performance depends on the sample size and tissue for DNA extraction in the discovery and target samples, the underlying genetic architecture and stability of DNA methylation changes and is also related to the period between an event and DNA methylation measurement. The MSs that were analyzed in our studies may more appropriately be considered as biomarkers of an exposure or a trait[83,84], rather than a predictor of a future condition.

Overall, these findings suggest that PGS and MS account for different aspects of the molecular biology underlying the traits and exposures of interest. Future studies are needed to investigate what kind of information is captured by GWASs and EWASs, and how this information can be used to improve multi-omics prediction of the traits and exposures. In addition, the nature of the common and non-common genes and GO terms could be investigated in more detail to gain more insight in the differential biological meaning of the EWAS and GWAS results. We suppose that for exposures polygenic scores should reflect the behavioral component, while methylation scores should reflect the effect of exposure as well.

Twins as subject of research

In this section of discussion, I would like to reflect on the role of twins who were the main participants in the various studies in this thesis.

Databases collected from twins can address twin specific questions and issues, and focus on research questions, which are relevant to the twin population. Data from twins can be included in a discovery study, using statistical methods to accommodate relatedness, or meta-analysed with other population-based results, as we did in the EWAS of left-handedness. The outcome from twin-singleton comparisons across a wide range of traits indicate that twins can be added to association studies of singletons for many traits and exposures, to increase sample sizes, which is crucial in genetic studies. One affirmation that twins can be included in population-based GWASs, was given by study of Beck et al.[85]. They reported robust positive genetic correlations, that do no differ from unity, between the meta-analysis results of twin birth weight and previously published GWAS results of singleton birth weight.

Millions of twins have participated in twin research all over the world, and their contributions have created a treasure trove of data[86]. Collaborative efforts pooled together the twin registries from different countries by many consortiums, such as NorTwinCan, IGEMS, GenomEUTwin, and CODATwins[86]. An excellent example of such collaborative work was realized by a group of twin researchers who had collected epigenetic data on twins. Combining resources from the largest world twin registers from the Netherlands, Australia, United Kingdom, and Finland, these researchers made a breakthrough discovery in the field of biology of monozygotic twinning and found an epigenetic signature that can detect if an individual is a part of the monozygotic twin pair[87]. This finding of an epigenetic signature of monozygotic twinning can contribute to the studies of the early embryonic development, influence of assisted reproduction technologies and infertility treatment, of the twinning split, the etiology of many congenital disorders, and the vanishing twin syndrome. This is an example of game-changing epigenetic research using twin data.

Research on twins and the outcomes of such studies may create specific ethical questions that do not arise in singleton studies, such as a necessity of participation of both twins, the use of information on heritable traits for individuals who have the same genome, or presenting information to twins when they are identified as discordant for traits or exposures. Researchers working in behavior genetics and genetic epidemiology are aware of such questions and also of psychological issues that may be unique to twins, but do not usually focus their research on ‘twin-specific’ questions. However, there are experienced twin researchers (who happen to be twins themselves), Nancy Segal and Joleen Greenwood, who pay specific attention to these issues, and recently published books on the twinning experience and the participation of twins in research[88,89].

Future directions

Evidence-based research benefits from a prior systematic review of findings related to the research question. The task of comprehensive monitoring of the field of research is increasingly time consuming given the growth of scientific literature and open access

datasets. The new technologies of automation are transforming the nature of research work, decreasing the burden of research routines, such as screening of the science literature. Automated procedures to screen for relevant literature, based on machine learning and natural language processing, helped us to systematize a huge body of literature, and demonstrated the reduction in workload. We were the first to use an early version of automated tool ASReview, which was developed by van de Schoot and colleagues at Utrecht University. Today, the ASReview is a complete software tool that can either be installed locally or on a server. ASReview was designed for alleviating the manual systematic reviewing process. It can however also be used to screen large amounts of any textual data as long as its properly formatted (https://asreview.nl). More automated tools employing machine learning and deep learning are needed to extract meaningful information from the articles and (epi)genetic summary statistics with reduced risk of bias, better visualization, and flexible systematization options.

My studies of early life characteristics and traits were focused on establishing associations in large datasets. The central problem in the interpretation of the association between prenatal exposure and offspring risk for certain disorders is to distinguish the degree to which the association is driven by a causal effect of prenatal factors or by confounding variables that predict both the risk of prenatal factors in the mother and outcome risk in her offspring[90]. There are several ways in which DNA methylation data, as a methylation score, can be used to study the relationship between early life factors and later outcomes. First, a methylation score can be used as a measured confounder. This approach is used in some studies, in which the *AHRR* methylation is used as surrogate variable to correct for smoking[55]. We exemplified the use of such surrogate variable features in our study of left-handedness in ALSPAC EWASs. Development of epigenetic studies could bridge the gap in the search for hidden confounding (unobserved factors) that influence both the risk factor and the outcome, as DNA methylation contains a signature of both, and of multiple confounders. Second, a methylation score could be included as a hypothesized causal variable, a substitute of an early life factor.

My epigenetic studies were mainly focused on associations between methylation and early life characteristics and did not consider the problem of causality. To study causality between DNA methylation signatures and scores and early life characteristics, further studies with other methodological approaches are needed. One approach is the discordant monozygotic twin-pair design, which can accommodate genetic and shared environmental confounding[91]. Twin research has a unique potential to contribute to the epigenomics revealing epigenetic profiles related to complex traits and disease through different twin models that are capable of disentangling the sources of phenotypic variance, the genetic and environmental influences on the epigenome, and to control for confounding in within pair designs. Another approach to study causality is Mendelian Randomization (MR), which is used to assess the causality of the modifiable exposure on an outcome by using instrumental genetic variants that are robustly associated with the exposure[92]. Given such genetic variants, this approach could be applied to study the association between left-handedness and breastfeeding[34,35], for example.

Epigenetic studies can contribute to the network of phenotypes, disorders, and disease genes that link known disorder-gene associations[93]. Integration of genome-wide

and epigenome-wide data will bring new insights to the genetic architecture of many complex traits, as is currently the case in the network-based modeling approaches in cancer research[94]. Development of multiple DNA methylation estimators of different complex traits, diseases, and exposures will contribute to the active translation of EWAS findings to the level of the individual's risk prediction[53,87,95–98]. DNA-methylation predictors developed on the basis of population research should demonstrate sufficient relevance before they are applied in individual diagnosis and research. Several DNA methylation estimators have been developed: tests for imprinted disorders[95], multiple DNA methylation age estimators or "epigenetic clocks"[97,99,100], lifestyle and biochemical traits[98], classifiers to determine smoking status[68], and epigenetic predictors of mortality[96]. Recently, a promising DNA-methylation-based estimator of monozygotic twinning was published[87]. All these advances contribute to the development of personalized genomics.

Throughout the period of my research, the field of personalized epigenomics displayed fast progress. Following the initial successes in prediction of body mass index based on stable epigenomic signatures[101], it was expected a decade ago that epigenomics would "get personal"[102]. Today several commercial epigenetic tests based on currently available DNA-methylation-based estimators are already available to consumers[103]. Much work needs to be to be done to prepare for the advent of low-cost epigenome reading, including solving issues regarding the information epigenomic data can provide about individuals[104]. Such issues need to be studied as carefully as the technological and biological discovery aspects of human epigenomics. I expect that one day the epigenetic and other omics biomarkers may be available for a monitoring of one's health, lifestyle changes, treatment, and healthy aging. I hope that these biomarkers will be accessible and interpretable for individuals and their medical doctors, and will provide early signs of disorders for prevention, differentiation diagnostics, and prompt intervention.

References

1. Joubert, B. R. *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *The American Journal of Human Genetics* **98**, 680–696 (2016).
2. Janssen, B. G. *et al.* Placental mitochondrial DNA and CYP1A1 gene methylation as molecular signatures for tobacco smoke exposure in pregnant women and the relevance for birth weight. *Journal of translational medicine* **15**, 5 (2017).
3. Van Otterdijk, S. D., Binder, A. M. & Michels, K. B. Locus-specific DNA methylation in the placenta is associated with levels of pro-inflammatory proteins in cord blood and they are both independently affected by maternal smoking during pregnancy. *Epigenetics* **12**, 875–885 (2017).
4. Banik, A. *et al.* Maternal Factors that Induce Epigenetic Changes Contribute to Neurological Disorders in Offspring. *Genes* **8**, (2017).
5. Cuellar-Partida, G. *et al.* Genome-wide association study identifies 48 common genetic variants associated with handedness. *Nature human behaviour* **5**, 59–70 (2021).
6. Van Dongen, J. *et al.* DNA methylation signatures of aggression and closely related constructs: A meta-analysis of epigenome-wide studies across the lifespan. *Molecular psychiatry* **26**, 2148–2162 (2021).
7. Van der Laan, C. *et al.* Continuity of genetic risk for aggressive behavior across the life-course. *Behavioral Genetics* (2021).
8. Nyirenda, M. J. & Byass, P. Pregnancy, programming, and predisposition. The Lancet. *Global health* **7**, e404–e405 (2019).
9. Banderali, G. *et al.* Short and long term health effects of parental tobacco smoking during pregnancy and lactation: a descriptive review. *Journal of translational medicine* **13**, 327 (2015).
10. Kurita, H. *et al.* Maternal alcohol consumption and risk of offspring with congenital malformation: the Japan Environment and Children's Study. *Pediatric Research* **90**, 479–486 (2021).
11. Roseboom, T. J. Epidemiological evidence for the developmental origins of health and disease: effects of prenatal undernutrition in humans. *Journal of Endocrinology* **242**, T135–T144 (2019).
12. Scholl, T. O. & Johnson, W. G. Folic acid: influence on the outcome of pregnancy. *The American journal of clinical nutrition* **71**, 1295S–303S (2000).
13. Mortensen, L. H. Socioeconomic inequality in birth weight and gestational age in Denmark 1996–2007: Using a family-based approach to explore alternative explanations. *Social Science & Medicine* **76**, 1–7 (2013).
14. Milsom, I. *et al.* Influence of maternal, obstetric and fetal risk factors on the prevalence of birth asphyxia at term in a Swedish urban population. *Acta obstetricia et gynecologica Scandinavica* **81**, 909–917 (2002).
15. Thorngren-Jerneck, K. & Herbst, A. Low 5-minute Apgar score: a population-based register study of 1 million term births. *Obstetrics and gynecology* **98**, 65–70 (2001).
16. Lamichhane, N. *et al.* Associations between maternal stress during pregnancy and offspring obesity risk later in life—A systematic literature review. *Obesity Reviews* **21**, e12951 (2020).
17. Victora, C. G. *et al.* Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *The Lancet* **387**, 475–490 (2016).
18. Barker, D. J. Fetal origins of coronary heart disease. *BMJ (Clinical research ed.)* **311**, 171–174 (1995).

19. Barker, D. J. In utero programming of chronic disease. *Clinical science (London, England : 1979)* **95**, 115–128 (1998).

20. Barker, D. J. P. Adult Consequences of Fetal Growth Restriction. *Clinical Obstetrics and Gynecology* **49**, 270–283 (2006).

21. Bateson, P. *et al.* Developmental plasticity and human health. *Nature* **430**, 419–421 (2004).

22. Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: a landscape takes shape. *Cell* **128**, 635–638 (2007).

23. Bouchard, L. *et al.* Leptin gene epigenetic adaptation to impaired glucose metabolism during pregnancy. *Diabetes care* **33**, 2436–2441 (2010).

24. Mamun, A. A., O'Callaghan, M. J., Williams, G. M. & Najman, J. M. Maternal smoking during pregnancy predicts adult offspring cardiovascular risk factors – evidence from a community-based large birth cohort study. *PloS one* **7**, e41106 (2012).

25. Sikdar, S. *et al.* Comparison of smoking-related DNA methylation between newborns from prenatal exposure and adults from personal smoking. *Epigenomics* **11**, 1487–1500 (2019).

26. Wiklund, P. *et al.* DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. *Clinical epigenetics* **11**, 97 (2019).

27. Gaunt, T. R. *et al.* Systematic identification of genetic influences on methylation across the human life course. *Genome Biology* **17**, 61 (2016).

28. Hanson, M. A. & Gluckman, P. D. Early developmental conditioning of later health and disease: physiology or pathophysiology? *Physiological reviews* **94**, 1027–1076 (2014).

29. Merjonen, P., Dolan, C. v., Bartels, M. & Boomsma, D. I. Does Breastfeeding Behavior Run in Families? Evidence From Twins, Their Sisters and Their Mothers in the Netherlands. *Twin Research and Human Genetics* **18**, 179–187 (2015).

30. Walker-Short, E. *et al.* Epigenome-Wide Association Study of Infant Feeding and DNA Methylation in Infancy and Childhood in a Population at Increased Risk for Type 1 Diabetes. *Nutrients* **13**, (2021).

31. Medland, S. E., Duffy, D. L., Wright, M. J., Geffen, G. M. & Martin, N. G. Handedness in Twins: Joint Analysis of Data From 35 Samples. *Twin Research and Human Genetics* **9**, 46–53 (2006).

32. Ocklenburg, S. et al. Epigenetic regulation of lateralized fetal spinal gene expression underlies hemispheric asymmetries. *eLife* **6**, (2017).

33. De Vries, J. I. P. *et al.* Fetal handedness and head position preference: A developmental study. *Developmental Psychobiology* **39**, 171–178 (2001).

34. Hujoel, P. P. Breastfeeding and handedness: a systematic review and meta-analysis of individual participant data. *Laterality* **24**, 582–599 (2019).

35. Denny, K. Breastfeeding predicts handedness. *Laterality* **17**, 361–368 (2012).

36. Van Rooij, J. *et al.* Evaluation of commonly used analysis strategies for epigenome- and transcriptome-wide association studies through replication of large-scale population studies. *Genome Biology* **20**, 235 (2019).

37. Willemsen, G. *et al.* The Netherlands Twin Register Biobank: A Resource for Genetic Epidemiological Studies. *Twin Research and Human Genetics* **13**, 231–245 (2010).

38. Pauwels, S. *et al.* The Influence of the Duration of Breastfeeding on the Infant's Metabolic Epigenome. *Nutrients* **11**, (2019).

39. Sherwood, W. B. *et al.* Duration of breastfeeding is associated with leptin (LEP) DNA methylation profiles and BMI in 10-year-old children. *Clinical Epigenetics* **11**, 128 (2019).

40. Obermann-Borst, S. A. *et al.* Duration of breastfeeding and gender are associated with methylation of the LEPTIN gene in very young children. *Pediatric Research* **74**, 344–349 (2013).

41. Soto-Ramírez, N. *et al.* The interaction of genetic variants and DNA methylation of the interleukin-4 receptor gene increase the risk of asthma at age 18 years. *Clinical Epigenetics* **5**, 1 (2013).

42. Hartwig, F. P., Loret de Mola, C., Davies, N. M., Victora, C. G. & Relton, C. L. Breastfeeding effects on DNA methylation in the offspring: A systematic literature review. *PLOS ONE* **12**, e0173070 (2017).

43. Naumova, O. Yu. *et al.* A Study of the Association between Breastfeeding and DNA Methylation in Peripheral Blood Cells of Infants. *Russian Journal of Genetics* **55**, 749–755 (2019).

44. Michels, K. B. *et al.* Recommendations for the design and analysis of epigenome-wide association studies. *Nature Methods* **10**, 949–955 (2013).

45. Mansell, G. *et al.* Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. *BMC Genomics* **20**, 366 (2019).

46. Suderman, M. *et al.* dmrff: identifying differentially methylated regions efficiently with power and control. *bioRxiv* 508556 (2018) doi:10.1101/508556.

47. Jaffe, A. E. *et al.* Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *International Journal of Epidemiology* **41**, 200–209 (2012).

48. Hüls, A. & Czamara, D. Methodological challenges in constructing DNA methylation risk scores. *Epigenetics* **15**, 1–11 (2020).

49. Katsanis, S. H. & Katsanis, N. Molecular genetic testing and the future of clinical genomics. *Nature Reviews Genetics* **14**, 415–426 (2013).

50. Wray, N. R., Goddard, M. E. & Visscher, P. M. Prediction of individual genetic risk of complex disease. *Current Opinion in Genetics & Development* **18**, 257–263 (2008).

51. Bitarello, B. D. & Mathieson, I. Polygenic Scores for Height in Admixed Populations. *G3 Genes|Genomes|Genetics* **10**, 4027–4036 (2020).

52. Lloyd-Jones, L. R. *et al.* Improved polygenic prediction by Bayesian multiple regression on summary statistics. *Nature Communications* **10**, 5086 (2019).

53. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics* **19**, 371–384 (2018).

54. Elliott, H. R. *et al.* Differences in smoking associated DNA methylation patterns in South Asians and Europeans. *Clinical epigenetics* **6**, 4 (2014).

55. Sugden, K. *et al.* Establishing a generalized polyepigenetic biomarker for tobacco smoking. *Translational psychiatry* **9**, 92 (2019).

56. Shah, S. *et al.* Improving Phenotypic Prediction by Combining Genetic and Epigenetic Associations. *The American Journal of Human Genetics* **97**, 75–85 (2015).

57. McCartney, D. L. *et al.* Epigenetic prediction of complex traits and death. *Genome biology* **19**, 136 (2018).

58. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**, 25–29 (2000).

59. Ensembl Variant Effect Predictor (VEP). https://grch37.ensembl.org/info/docs/tools/vep/index.html

60. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, (2016).

61. Kinsella, R. J. *et al.* Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database* **2011**, (2011).

62. Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P. D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research* **47**, D419–D426 (2019).

63. Carbon, S. *et al.* The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Research* **49**, D325–D334 (2021).

64. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics 2000 25:1* **25**, 25–29 (2000).

65. Joehanes, R. *et al.* Epigenetic Signatures of Cigarette Smoking. *Circulation. Cardiovascular genetics* **9**, 436–447 (2016).

66. Küpers, L. K. *et al.* DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *International journal of epidemiology* **44**, 1224–1237 (2015).

67. Richmond, R. C., Suderman, M., Langdon, R., Relton, C. L. & Davey Smith, G. DNA methylation as a marker for prenatal smoke exposure in adults. *International journal of epidemiology* **47**, 1120–1130 (2018).

68. Bollepalli, S., Korhonen, T., Kaprio, J., Anders, S. & Ollikainen, M. EpiSmokEr: a robust classifier to determine smoking status from DNA methylation data. *Epigenomics* **11**, 1469–1486 (2019).

69. Zhang, Y. *et al.* Smoking-Associated DNA Methylation Biomarkers and Their Predictive Value for All-Cause and Cardiovascular Mortality. *Environmental Health Perspectives* **124**, 67–74 (2016).

70. Reese, S. E. *et al.* DNA Methylation Score as a Biomarker in Newborns for Sustained *Maternal Smoking during Pregnancy* **125**, 760–766 (2017).

71. Gao, X., Thomsen, H., Zhang, Y., Breitling, L. P. & Brenner, H. The impact of methylation quantitative trait loci (mQTLs) on active smoking-related DNA methylation changes. *Clinical Epigenetics* **9**, 87 (2017).

72. Madden, R. A. *et al.* Birth weight associations with DNA methylation differences in an adult population. *Epigenetics* **16**, 783–796 (2021).

73. Tan, Q. *et al.* Epigenetic signature of preterm birth in adult twins. *Clinical Epigenetics* **10**, 87 (2018).

74. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Human molecular genetics* **27**, 3641–3649 (2018).

75. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature genetics* **47**, 1228–1235 (2015).

76. Küpers, L. K. *et al.* Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight 10, 1893 (2019).

77. Ip, H. F. *et al.* Genetic association study of childhood aggression across raters, instruments and age. *Transl Psychiatry* **11**, 413 (2021).

78. Ocklenburg, S. *et al.* Polygenic scores for handedness and their association with asymmetries in brain structure. *Brain structure & function* (2021) doi:10.1007/s00429-021-02335-3.

79. De Kovel, C. G. F. & Francks, C. The molecular genetics of hand preference revisited. *Scientific Reports* **9**, 5986 (2019).

80. Wray, N. R. *et al.* Research Review: Polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry* **55**, 1068–1087 (2014).

81. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *American journal of human genetics* **101**, 5–22 (2017).

82. Wittgenstein, L. (Ludwig J. J. & Anscombe, G. E. M. (Gertrude E. M. *Philosophical investigations* (Blackwell, 1968).

83. García-Giménez, J. L. *et al.* Epigenetic biomarkers: Current strategies and future challenges for their use in the clinical laboratory. *Critical reviews in clinical laboratory sciences* **54**, 529–550 (2017).

84. Reed, Z. E., Suderman, M. J., Relton, C. L., Davis, O. S. P. & Hemani, G. The association of DNA methylation with body mass index: distinguishing between predictors and biomarkers. *Clinical epigenetics* **12**, 50 (2020).

85. Beck, J. J. *et al.* Genetic meta-analysis of twin birth weight shows high genetic correlation with singleton birth weight. *Human Molecular Genetics* (2021) doi:10.1093/hmg/ddab121.

86. Hur, Y., Odintsova, V., Ordonana, J., Silventoinen, K. & Willemsen, G. Twin Registries Worldwide. in *Twin and Family Studies of Epigenetics* (eds. Harris, J., Segal, N., Tarnoki, A. D. & Tarnoki, D. L.) (Elsevier, 2022).

87. Van Dongen, J. *et al.* Identical twins carry a persistent epigenetic signature of early genome programming. *Nature Communications* **12**, 5618 (2021).

88. Segal, N. L. 1951-T. A.-T. T.-. Deliberately divided: inside the controversial study of twins and triplets adopted apart (2021).

89. Greenwood, J. Loucks. *Identical Twins: adult reflections on the twinship experience* (LEXINGTON BOOKS, 2018).

90. Kendler, K. S. Maternal Bacterial Infection During Pregnancy: A Potential Causal Risk Factor for Psychosis in Offspring. *American Journal of Psychiatry* **177**, 14–16 (2020).

91. Vitaro, F., Brendgen, M. & Arseneault, L. The discordant MZ-twin method: One step closer to the holy grail of causality. *International Journal of Behavioral Development* **33**, 376–382 (2009).

92. Evans, D. M. & Davey Smith, G. Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality. *Annual Review of Genomics and Human Genetics* **16**, 327–350 (2015).

93. Goh, K.-I. *et al.* The human disease network. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 8685–90 (2007).

94. Sanchez, R. & Mackenzie, S. A. Integrative Network Analysis of Differentially Methylated and Expressed Genes for Biomarker Identification in Leukemia. *Scientific Reports* **10**, 2123 (2020).

95. Aref-Eshghi, E. *et al.* Evaluation of DNA Methylation Episignatures for Diagnosis and Phenotype Correlations in 42 Mendelian Neurodevelopmental Disorders. *American journal of human genetics* **106**, 356–370 (2020).

96. Lu, A. T. *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303–327 (2019).

97. McEwen, L. M. *et al.* The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 23329–23335 (2020).

98. Hillary, R. F. & Marioni, R. E. MethylDetectR: a software for methylation-based health profiling. *Wellcome Open Research* **5**, 283 (2021).

99. Horvath, S. DNA methylation age of human tissues and cell types. *Genome biology* **14**, R115 (2013).

100. Hannum, G. *et al.* Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell* **49**, 359–367 (2013).

101. Feinberg, A. P. *et al.* Personalized Epigenomic Signatures That Are Stable Over Time and Covary with Body Mass Index. *Science Translational Medicine* **2**, 49ra67 (2010).

102. Flintoft, L. Epigenomics gets personal. *Nature Reviews Genetics* **11**, 747–747 (2010).

103. Dupras, C., Beauchamp, E. & Joly, Y. Selling direct-to-consumer epigenetic tests: are we ready? *Nature reviews. Genetics* **21**, 335–336 (2020).

104. Alex, K. & Winkler, E. C. Is Dupras and Bunnik's Framework for Assessing Privacy Risks in Multi-Omic Research and Databases Still Too Exceptionalist? *The American Journal of Bioethics* **21**, 80–82 (2021).

Supplemental Table 1.
Follow-up of PGS and MS prediction: Overlaps between GWAS and EWAS gene sets

gene	chr	start	end
Smoking			
no overlaps			
Prenatal maternal smoking			
MAML3	4	140637907	141075338
Body mass index			
MAD1L1	7	1855429	2272878
CALN1	7	71244476	71912136
PCGF5	10	92979908	93044088
CPT1A	11	68522088	68611878
CACNA1C	12	2079952	2802108
CHFR	12	133398773	133532890
BTBD7	14	93703896	93799438
SLCO3A1	15	92396925	92715665
AXIN1	16	337440	402673

Supplemental Table 2.
Follow-up of PGS and MS prediction: Overlaps between GWAS and EWAS GO biological process terms for body mass index

Biological process	Gene Ontology code	*Fold Enrich-ment*	GWAS *P*-value	FDR	*Fold Enrich-ment*	GWAS *P*-value	FDR
nervous system development	GO:0007399	1.93	1.2E-12	1.9E-08	1.75	1.9E-04	4.2E-02
positive regulation of transcription by RNA polymerase II	GO:0045944	1.69	1.2E-04	1.4E-02	2.41	2.0E-06	1.0E-03
positive regulation of nucleic acid-templated transcription	GO:1903508	1.64	2.9E-05	4.7E-03	2.08	9.7E-06	3.8E-03
positive regulation of transcription, DNA-templated	GO:0045893	1.64	2.9E-05	4.6E-03	2.08	9.7E-06	3.7E-03
positive regulation of RNA biosynthetic process	GO:1902680	1.63	2.9E-05	4.6E-03	2.08	9.9E-06	3.7E-03
positive regulation of RNA metabolic process	GO:0051254	1.63	1.5E-05	3.0E-03	2.16	1.1E-06	6.0E-04

Supplemental Table 2. (continued)

Biological process	Gene Ontology code	*Fold Enrich-ment*	GWAS *P*-value	FDR	*Fold Enrich-ment*	GWAS *P*-value	FDR
positive regulation of nucleobase-containing compound metabolic process	GO:0045935	1.63	4.7E-06	1.3E-03	2.11	9.7E-07	5.6E-04
positive regulation of macromolecule biosynthetic process	GO:0010557	1.59	2.5E-05	4.3E-03	1.9	5.4E-05	1.5E-02
regulation of signaling	GO:0023051	1.59	5.6E-09	1.8E-05	1.93	6.2E-09	9.8E-06
regulation of cell communication	GO:0010646	1.59	7.4E-09	1.9E-05	1.94	5.6E-09	1.3E-05
positive regulation of biosynthetic process	GO:0009891	1.56	2.5E-05	4.3E-03	1.85	6.2E-05	1.7E-02
positive regulation of cellular biosynthetic process	GO:0031328	1.56	3.9E-05	5.9E-03	1.85	8.5E-05	2.2E-02
positive regulation of cellular metabolic process	GO:0031325	1.54	1.1E-07	1.2E-04	1.95	7.3E-09	9.6E-06
regulation of signal transduction	GO:0009966	1.54	7.3E-07	4.2E-04	1.88	3.5E-07	2.5E-04
positive regulation of nitrogen compound metabolic process	GO:0051173	1.52	8.6E-07	4.8E-04	1.99	5.1E-09	1.3E-05
positive regulation of metabolic process	GO:0009893	1.51	4.0E-08	7.9E-05	1.85	1.2E-08	1.3E-05
animal organ development	GO:0048513	1.51	8.7E-07	4.7E-04	1.73	8.0E-06	3.3E-03
system development	GO:0048731	1.51	3.6E-09	1.4E-05	1.67	8.1E-07	4.9E-04
positive regulation of macromolecule metabolic process	GO:0010604	1.49	5.3E-07	3.4E-04	1.91	6.3E-09	9.0E-06
multicellular organism development	GO:0007275	1.47	7.5E-09	1.7E-05	1.66	2.3E-07	1.7E-04
regulation of multicellular organismal process	GO:0051239	1.45	8.0E-05	1.0E-02	1.78	2.2E-05	7.4E-03
regulation of response to stimulus	GO:0048583	1.42	2.9E-06	9.7E-04	1.76	1.3E-07	1.1E-04
regulation of macromolecule biosynthetic process	GO:0010556	1.42	2.3E-06	8.8E-04	1.56	5.2E-05	1.5E-02
regulation of cellular macromolecule biosynthetic process	GO:2000112	1.41	3.6E-06	1.1E-03	1.57	4.8E-05	1.4E-02
anatomical structure development	GO:0048856	1.41	5.5E-08	9.6E-05	1.6	4.0E-07	2.6E-04
regulation of biosynthetic process	GO:0009889	1.4	2.8E-06	9.9E-04	1.57	1.8E-05	6.2E-03
regulation of cellular biosynthetic process	GO:0031326	1.4	4.1E-06	1.2E-03	1.58	1.7E-05	5.9E-03

Supplemental Table 2. (continued)

Biological process	Gene Ontology code	*Fold Enrich-ment*	GWAS *P*-value	FDR	*Fold Enrich-ment*	GWAS *P*-value	FDR
regulation of transcription by RNA polymerase II	GO:0006357	1.39	1.9E-04	4.7E-02	1.77	2.9E-05	9.1E-03
positive regulation of cellular process	GO:0048522	1.38	5.2E-04	9.0E-05	1.7	2.1E-10	1.6E-06
regulation of RNA metabolic process	GO:0051252	1.36	5.8E-08	8.8E-03	1.66	5.1E-06	2.4E-03
regulation of transcription, DNA-templated	GO:0006355	1.36	6.7E-05	2.1E-02	1.63	3.9E-05	1.2E-02
regulation of nucleic acid-templated transcription	GO:1903506	1.36	1.9E-04	2.1E-02	1.63	3.9E-05	1.2E-02
regulation of RNA biosynthetic process	GO:2001141	1.36	1.9E-04	2.1E-02	1.62	3.9E-05	1.2E-02
regulation of nucleobase-containing compound metabolic process	GO:0019219	1.35	6.4E-05	8.4E-03	1.64	3.8E-06	1.8E-03
negative regulation of cellular process	GO:0048523	1.35	6.9E-06	1.9E-03	1.73	3.3E-09	1.1E-05
negative regulation of biological process	GO:0048519	1.34	2.0E-06	8.2E-04	1.65	1.4E-08	1.5E-05
developmental process	GO:0032502	1.34	6.8E-07	4.1E-04	1.53	1.5E-06	8.0E-04
regulation of gene expression	GO:0010468	1.34	1.6E-05	3.1E-03	1.51	2.3E-05	7.7E-03
positive regulation of biological process	GO:0048518	1.33	5.1E-07	3.5E-04	1.66	1.5E-10	2.4E-06
regulation of cellular metabolic process	GO:0031323	1.31	3.4E-06	1.1E-03	1.6	1.0E-08	1.2E-05
regulation of nitrogen compound metabolic process	GO:0051171	1.3	1.9E-05	3.3E-03	1.61	2.8E-08	2.6E-05
regulation of macromolecule metabolic process	:0060255	1.29	7.5E-06	1.9E-03	1.6	5.6E-09	1.1E-05
regulation of metabolic process	GO:0019222	1.29	2.7E-06	9.9E-04	1.59	9.6E-10	5.0E-06
regulation of primary metabolic process	GO:0080090	1.28	4.9E-05	6.9E-03	1.65	1.8E-09	7.2E-06
cellular response to stimulus	GO:0051716	1.25	7.0E-05	9.2E-03	1.38	9.0E-05	2.3E-02
multicellular organismal process	GO:0032501	1.25	4.7E-05	6.6E-03	1.35	2.3E-04	4.8E-02
regulation of cellular process	GO:0050794	1.19	6.6E-07	4.1E-04	1.32	3.4E-08	3.0E-05
regulation of biological process	GO:0050789	1.18	3.0E-07	2.2E-04	1.31	1.5E-08	1.4E-05
biological regulation	GO:0065007	1.16	1.6E-06	7.2E-04	1.25	1.5E-06	7.7E-04
biological process	GO:0008150	1.06	9.2E-05	1.2E-02	1.09	1.6E-04	3.7E-02

Appendices

LIST OF AUTHOR'S CONTRIBUTIONS IN THESIS CHAPTERS

Chapter 2

Published as: Odintsova, V. V., Dolan, C. V., van Beijsterveldt, C., de Zeeuw, E. L., van Dongen, J., & Boomsma, D. I. (2019). Pre- and Perinatal Characteristics Associated with Apgar Scores in a Review and in a New Study of Dutch Twins. *Twin research and human genetics.* 22(3), 164–176.

My contribution to Chapter 2 comprised the selection and review of literature on Apgar scores in singletons and twins. With Conor Dolan and Dorret Boomsma, I defined the concept of the paper, selected early life characteristics to be included in the analysis in the NTR. Toos van Beijsterveldt and Eveline de Zeeuw collected, did the quality control (QC), and preprocessed data in the NTR. I wrote the data request and analysis plan, and prepared the data for analysis together with Toos van Beijsterveldt. I performed all analyses under supervision of Conor Dolan. I created tables and figures with input of Eveline de Zeeuw. I wrote the paper, processed edits from co-authors, and was responsible for submission and revision. All co-authors actively participated in editing and reviewing the manuscript, and I processed all comments and feedback prior to submission of the manuscript.

Chapter 3

Published as: Odintsova, V.V., Saifitdinova, A.F., Naumova, O.Y. (2018). Maternal smoking and DNA methylation abnormalities in children at early developmental stages. *Akusherstvo i Ginekologiya (Russian Federation),* 9, pp. 5-12.

For Chapter 3, I received funding from Russian Foundation of Basic Research. I performed the literature search and selected epigenome-wide association studies of prenatal maternal smoking to be included in the review. I contributed to the visualization performed by Alsu Saifitdinova. I wrote the first draft of the manuscript, that was substantially elaborated together with Oxana Naumova and Alsu Saifitdinova. I edited the last version of the manuscript and submitted it in Russian. I translated the paper into English. The text of the Chapter was reviewed by Conor Dolan, Dorret Boomsma, and Jenny van Dongen.

Chapter 4

Published as: Odintsova, V. V., Hagenbeek, F. A., Suderman, M., Caramaschi, D., van Beijsterveldt, C., Kallsen, N. A., Ehli, E. A., Davies, G. E., Sukhikh, G. T., Fanos, V., Relton, C., Bartels, M., Boomsma, D. I., & van Dongen, J. (2019). DNA Methylation Signatures of Breastfeeding in Buccal Cells Collected in Mid-Childhood. *Nutrients,* 11(11), 2804.

For Chapter 4, I obtained funding from Amsterdam Public Health (APH) for a working visit to Bristol University. The two research groups (the NTR and ALSPAC) designed the study with discovery study in NTR and replication in ALSPAC. I made a literature review on effects of breastfeeding with inputs from Gennady Sukhikh and Vassilios Fanos. I made a literature review on epigenome-wide association studies of breastfeeding with input from Doretta Caramaschi on an unpublished study of breastfeeding associated DNA methylation. I analyzed DNA methylation data in NTR generated by the Human Genotyping facility of ErasmusMC, the Netherlands, and Noah Kallsen, Gareth Davies and Erik Ehli at Avera Institute for Human Genetics. Dorret Boomsma, Jenny van Dongen, Fiona Hagenbeek and Meike Bartels supervised the NTR-ACTION Biobank sample resource and data. I performed analyses in NTR under supervision of Jenny van Dongen and coordinated with Matthew Suderman and Doretta Caramaschi on replication of findings in ALSPAC. I created all visualizations, wrote a first draft of manuscript, joined edits from all co-authors, and was responsible for submission and revision.

Chapter 5

Published as: Odintsova V., Dolan CV, Beijstervaldt T., Ligthart L., Willemsen G., de Geus E.J.C., van Dongen J., Boomsma DI. Handedness and 23 early life characteristics: a review and a study in 37,495 Dutch twins (submitted).

My contribution to Chapter 5 comprised the elaboration of the study design together with Conor Dolan and Dorret Boomsma. I made a literature review of the studies on association of handedness and early life factors, and systemized it. Together with Toos van Beijsterveldt and Lannie Ligthart, I selected the early life characteristics available in the NTR, and prepared a data request. I performed all analyses consulting with Conor Dolan on statistical analysis. I wrote the manuscript with active participation of all co-authors in editing and reviewing the manuscript. I joined the edits, and submitted the manuscript.

Chapter 6

Published as: Odintsova V.*, Suderman M.*, Hagenbeek F., Caramaschi D., Hottenga J.J., Pool R., BIOS consortium, Dolan C.V., Ligthart L., van Beijsterveldt C.E.M., Willemsen G., de Geus E.J.C., Beck J.J., Ehli E.A., Cuellar-Partida G., Evans D.M., Medland S.E., Relton C.L., Boomsma D.I., van Dongen J. (2022) DNA methylation in peripheral tissues and left-handedness. *Scientific Reports,* 2022. 12(1): 5606.

In Chapter 6, I shared first authorship with Matthew Suderman (MRC Integrative Unit of University of Bristol, ALSPAC). This was the second paper of our collaboration. Together with Jenny van Dongen, Dorret Boomsma and Caroline Relton, we developed the study design. Matthew Suderman performed analyses in ALSPAC parents and offspring, and I performed analyses in the NTR adults and children, and did a meta-analysis. Jenny van Dongen supervised the methodology of the epigenome-wide association study and polygenic and DNA methylation score prediction. Matthew Suderman supervised the differentially regions analysis. Gabriel Cuellar-Partida and Sara Medland provided summary statistics from recent genome-wide association study of handedness. Rene Pool and Jouke-Jan Hottenga calculated polygenic scores for left-handedness based on this summary statistics. Fiona Hagenbeek, Gonneke Willemsen, Eco de Geus, Jenny van Dongen, and Dorret Boomsma created NTR-Action and NTR Biobank sample resources and data. The BIOS consortium and Avera Institute of Human Genetics (Jeffrey Beck and Erik Ehli) generated DNA methylation data for NTR. Toos van Beijsterveldt and Lannie Ligthart collected, cleaned, and preprocessed handedness data in the NTR. I prepared the first draft of the manuscript. All co-authors actively participated in editing and reviewing the manuscript, and Jenny van Dongen, Matthew Suderman and I collaborated on processing all comments and feedback prior to submission. The manuscript passed first round of revision at Scientific Reports.

Chapter 7

Published as: Odintsova V.V.*, Roetman P.J*, Pool R., Ip H. F., Van der Laan C.M., Tona D.K., Vermeiren R.R.J.M., Boomsma D. I. (2019). Genomics of human aggression: current state of genome-wide studies and an automated systematic review tool. *Psychiatric Genetics.* Oct;29(5):170-190.

The study was updated for publication as Odintsova VV, Hagenbeek FA, van der Laan CM, van de Weijer S., Boomsma DI. Genetics and epigenetics of human aggression. In: Handbook of Clinical Neurology. Brain and Crime (Eds. Swaab J.T.). Elsevier (2022) (in press).

In Chapter 7, I shared the first authorship with Peter Roetman. Together we performed manual selection of literature for a systematic review. In parallel, I performed selection of literature using machine-learning-based algorithms under supervision of developers from Utrecht University (Rens van de Schoot) and with active participance of Rene Pool. Together with Rene Pool, we compared traditional and automated approaches to literature selection. Together with Peter Roetman, we wrote a draft of paper. Together with Ip Hill and Camiel van der Laan, we extracted genetic variants associated with aggression-related traits, made calculations for each chromosome, and created visualizations. I created the tables, integrative

figure of behavioral genetic and molecular studies of aggression, joined edits from all co-authors, and was responsible for submission and revision.

I updated the study in 2021, and performed selection of literature and review of epigenetic studies, wrote a section “Epigenome-wide association studies”. Camiel van der Laan updated the genome-wide association studies review and visualization of number of genetic variants associated with aggression-related traits in overviewed GWASs. Fiona Hagenbeek made an addition on other omics domains. Dorret Boomsma and Steve van de Weijer revised all the text. I made visualizations, processed edits, updates and an additional section from co-authors, and was responsible for submission.

Chapter 8

Published as: Odintsova, V. V.*, Rebattu, V.*, Hagenbeek, F. A., Pool, R., Beck, J. J., Ehli, E. A., van Beijsterveldt, C., Ligthart, L., Willemsen, G., de Geus, E., Hottenga, J. J., Boomsma, D. I., & van Dongen, J. (2021) Predicting Complex Traits and Exposures From Polygenic Scores and Blood and Buccal DNA Methylation Profiles. *Frontiers in Psychiatry,* 12:688464.

In Chapter 7, I shared the first authorship with Valerie Rebattu who gathered and preprocessed summary statistics from discovery studies used for calculation of DNA methylation scores. I calculated blood and buccal DNA methylation scores for traits and exposures. Rene Pool and Jouke-Jan Hottenga calculated polygenic scores for these traits and exposures. I performed prediction of traits and exposures based on DNA methylation scores and polygenic scores under supervision of Jenny van Dongen. I compared our results with other reports on polygenic and DNA methylation scores in the literature and created visualizations of results. Toos van Beijsterveldt and Lannie Ligthart collected, cleaned, and preprocessed phenotype data in the NTR. Fiona Hagenbeek, Gonneke Willemsen, Eco de Geus, Jenny van Dongen, and Dorret Boomsma created NTR-Action and NTR Biobank sample resources and data. Avera Institute of Human Genetics (Jeffrey Beck and Erik Ehli) generated DNA methylation data for NTR. I wrote the first draft of the manuscript. All co-authors contributed to the editing and reviewing of the manuscript.

Chapter 9

Published as: Odintsova, V. V., Willemsen, G., Dolan, C. V., Hottenga, J. J., Martin, N. G., Slagboom, P. E., Ordoñana, J. R., & Boomsma, D. I. (2018). Establishing a twin register: an invaluable resource for (behavior) genetic, epidemiological, biomarker and «omics» studies. *Twin Research and Human Genetics*, (3), 1-14.

My contribution to the Chapter 9 comprised conceptualization of an idea that arose of the ISTS conference in Madrid together with Juan Ordoñana, Gonneke Willemsen and Dorret Boomsma. I integrated the materials from all co-authors, who contributed to the manuscript and wrote the first draft of the paper with Juan Ordoñana. Jouke-Jan Hottenga analyzed data for zygosity determination, the effect of zygosity misclassification on the heritability estimates, and made visualizations for this part of the manuscript. Conor Dolan carried out simulation studies of batch effects and family clustering, described in the Appendix. All co-authors actively participated in writing and editing. I was responsible for submission and revision.

Chapter 10

Published as: Willemsen, G., Odintsova, V., de Geus, E. & Boomsma, DI. (2021). Twin-singleton comparisons across multiple domains of life (chapter 4). In: Twin and Higher-order Pregnancies (eds. Khalil, A., Lewi, L. & Lopriore, E.) (Springer International Publishing), pp. 51-71.

In Chapter 10, I performed the review of literature on congenital disorders, morbidity and mortality in twins, wrote a draft of a part of the Chapter, and made edits of the final version. All co-authors actively participated in editing and reviewing the manuscript.

Chapter 11

Published as: Odintsova V. (2019). Book review: Identical Twins: Adult Reflections on the Twinship Experience - Joleen Loucks Greenwood. Twin Research and Human Genetics, 22(3), 199-199. Odintsova, V. (2021). Deliberately Divided: Inside the Controversial Study of Twins and Triplets Adopted Apart - Nancy Segal. *Twin Research and Human Genetics*, 1-2.

In Chapter 11, I wrote two reviews on recently published books, that were published in the journal Twin Research and Human Genetics with 2-year interval. The book "Identical Twins: Adult Reflections on the Twinship Experience" of Joleen Loucks Greenwood is published by Lexington books (154 pages). The book "Deliberately Divided: Inside the Controversial Study of Twins and Triplets Adopted Apart" is published by Rowman & Littlefield Publishers (507 pages). For the last review I analysed the documentaries "The Twinning Reaction" and "Three Identical Strangers" and book of Neubauer (1990).

Not part of this thesis

van Dongen, J., Gordon, S. D., McRae, A. F., Odintsova, V. V., Mbarek, H., Breeze, C. E., Sugden, K., Lundgren, S., Castillo-Fernandez, J. E., Hannon, E., Moffitt, T. E., Hagenbeek, F. A., van Beijsterveldt, C., Jan Hottenga, J., Tsai, P. C., BIOS Consortium, Genetics of DNA Methylation Consortium, Min, J. L., Hemani, G., Ehli, E. A., … Boomsma, D. I. (2021). Identical twins carry a persistent epigenetic signature of early genome programming. *Nature communications*, 12(1), 5618.

van Dongen J, Gordon SD, Odintsova VV, et al. (2021). Examining the Vanishing Twin Hypothesis of Neural Tube Defects: Application of an Epigenetic Predictor for Monozygotic Twinning. *Twin Research and Human Genetics,* 24(3):155-159.

van Dongen J, Odintsova V., Boomsma D. (2021). Chapter 3 "Discordant monozygotic twin studies of epigenetic mechanisms in mental health". In: Translational Epigenetics. *Twin and Family Studies of Epigenetics* (eds. Shuai L., Hopper J.L.). (Elsevier), pp.43-66.

Hur Y., Odintsova V., Ordonana J., Silventoinen K., Willemsen G. (2022). Chapter 3 "Twin Registries Worldwide". In: *Twin and Family Studies of Epigenetics* (eds. Harris J., Segal N., Tarnoki A.D., Tarnoki D.L.). (Elsevier)

Boomsma DI, van Beijsterveldt TCEM, Odintsova VV, Neale MC, Dolan CV. (2021). Genetically Informed Regression Analysis: Application to Aggression Prediction by Inattention and Hyperactivity in Children and Adults. *Behavior Genetics.* 51(3):250-263.

Rajula H.S.R., Odintsova V., Manchia M., Fanos V. (2019). Overview of federated facility to harmonize, analyze and management of missing data in cohorts. *Applied sciences*. 9(19):4103.

Naumova OY, Rychkov SY, Kornilov SA, Odintsova V.V., Anikina V.O., Solodunova M.Yu., Arintcina I.A., Zhukova M.A., Ovchinnikova I.V., Burenkova O.V., Zhukova O.V., Muhamedrahimov R.J., Grigorenko E.L. (2019). Effects of early social deprivation on epigenetic statuses and adaptive behavior of young children: A study based on a cohort of institutionalized infants and toddlers. *PLoS One,* 14(3):e0214285.

ABOUT THE AUTHOR

Veronika Odintsova was born in June 28, 1979 in Kazan, Russia. The daughter of Alfiya and Victor, both medical doctors, she studied Medicine at Kazan Medical State University where she was awarded her diploma *cum laude.* She continued her specialization in Healthcare and Public Health at the First Moscow Medical State University, and defended her first PhD thesis *"Project management approach in the field of mother and child health"* in 2008. After that she was involved in different research projects in epidemiology and healthcare at Saint-Petersburg State University, Federal Research Institute of Health Organization and Informatics, Moscow Department of Healthcare, and Kulakov National Research Center for Obstetrics, Gynecology and Perinatology. In addition to her education, she followed the specialization program in Psychology at Saint Petersburg State University and performed the graduate work *"Emotional well-being of patients of different clinical groups"* in 2011, graduating *cum laude.*

Supervised by principal investigator professor Elena Grigorenko and Oxana Naumova of Houston University, the USA, Veronika started epigenetic studies in 2014 within a large project *"Impact of Early Deprivation on Bio-behavioral Indicators of Child Development"*. In this project which utilized an adoption study design, she generated medical data and biosample collection from more than 200 children raised in institutional care and families. She subsequently became interested in twin study design and with the aim of developing her knowledge in Epigenetic Bioinformatics, Veronika started doctoral training at Vrije University of Amsterdam. There she joined the research group of prof. Dorret Boomsma and Jenny van Dongen, who work with the Netherlands Twin Register. Prof. Conor Dolan supervised her training in advanced statistics using twin design.

In 2019 Veronika was awarded an Amsterdam Public Health Institute grant, which gave her the opportunity to work as a visiting researcher at the MRC Integrative Epidemiological Unit of Bristol University, an expert centre for genetic and epigenetic epidemiology. In the last year of her PhD she won the first prize in vlogging competition of Amsterdam Public Health and received a grant support from Amsterdam Reproduction & Development to create videos for the dissemination of research findings.

PHD PORTFOLIO

PhD training

	Institute	Year	ECTs
General courses			
Advanced Academic Writing for PhD Researchers	Vrije University Amsterdam	2018	3.0
Grant writing and science communication	Vrije University Amsterdam	2020	6.0
Research Integrity - Biomedical sciences	Epigeum, Oxford University Press	2021	1.0
Statistical programming in R and Python	Vrije University Amsterdam	2020	6.0
Molecular Biology and Behavioral Genetics courses			
Behavioral Genetics	Vrije University Amsterdam	2018	6.0
Complex trait genetics	Vrije University Amsterdam	2018	6.0
Introduction to omics	Vrije University Amsterdam	2018	6.0
International Statistical Genetics Workshop “Twin Methodology”	Colorado University, Boulder	2018	2.0
Epigenomics and sequencing in behaviour and health	Vrije University Amsterdam	2019	6.0
Gene Finding: Genome-Wide Association Studies and Beyond	Vrije University Amsterdam	2020	6.0
Nature of Nurture: gene-environment correlation and interaction	Vrije University Amsterdam	2020	6.0
International Statistical Genetics Workshop “Structural Equation Modeling”	Colorado University, Boulder	2020	2.0
Mendelian Randomization in Medical Research	London Imperial College	2021	2.0
Enrichment analysis	Swiss Institute of Bioinformatics	2021	2.0
International conferences and meetings			
Joint Congress on Twin Pregnancy & Congress of the International Society Twin Studies (oral presentation)	Madrid, Spain	2017	2.0
48th Behavioral Genetics Association Meeting (poster presentation)	Boston, USA	2018	1.0
EU ACTION Project workshop (oral presentation)	Sardinia, Italy	2019	2.0
50th Behavioral Genetics Association Meeting (oral presentation)	online	2020	2.0
International Statistical Genetics Workshop (poster presentation)	Colorado University, Boulder, USA	2020	1.0
51th Behavioral Genetics Association Meeting (oral presentation)	online	2021	2.0
19th International Congress on Twin Studies (oral presentation)	online	2021	2.0
The First Advances in Precision Medicine (poster presentation)	Doha, Qatar	2021	2.0

Teaching activities

		Year	ECTs
Supervision bachelor and master thesis			
Supervision of research Internship I *(Valerie Rebattu)*	Vrije University Amsterdam	2020	3.0
Supervision of bachelor thesis *(Karlijn de Wit)*	Vrije University Amsterdam	2020	3.0
Teaching			
Research master course “Complex Trait Genetics” (assistant)	Vrije University Amsterdam	2020	1.0
R practical on EWAS in VU Summer School “Nature vs Nurture” (teacher)	Vrije University Amsterdam	2021	0.25
Research master course “Epigenomics and Sequencing” (assistance)	Vrije University Amsterdam	2021	1.0
Research master course “Complex Trait Genetics” (teacher)	Vrije University Amsterdam	2022	1.0
Measurement and Diagnostics for 2nd year bachelor students (tutorship, reviewing and grading assignments)	Vrije University Amsterdam	2021	4.0

Membership

Amsterdam Reproduction and Development (AR&D)	https://www.amsterdamumc.org/en/research/institutes/amsterdam-reproduction-development.htm
Amsterdam Public Health (Methodology Research Program)	https://www.amsterdamumc.org/en/research/institutes/amsterdam-public-health.htm
International Society of Twin Studies (ISTS)	https://twinstudies.org
Behavioral Genetics Association	http://bga.org

Fellowship

	Institute	Year	ECTs
Three-month fellowship at MRC Integrative Epidemiological Unit at Bristol University, UK, meta-analysis of NTR and ALSPAC cohorts	Bristol University	2019	6.0

Grants and awards

	Year
Russian Foundation for Basic Research grant for research project “Biosocial Research of Early Child Development and Epigenetic Status: Effects of Mothers’ Smoking and Breastfeeding on the Child DNA Methylation”	2018
Amsterdam Public Health, travel grant for working visit to Bristol University	2019
Pfizer-VU 1st Prize for brainstorming competition on “Real World Data collection and analysis to optimize treatment outcomes”	2021
Van der Gaag Grant of the Royal Netherlands Academy of Arts and Sciences (Koninklijke Nederlandse Akademie van Wetenschappen, KNAW) for working visit to Avera Institute of Human Genetics (USA) for research project “An epigenetic signature of monozygotic twinning: next steps”	2022

Dissemination projects

		Year
Article in Amsterdam Reproduction&Development annual report https://mcusercontent.com/530f953acd5069769d7066ae4/files/2310e3cb-a6bf-9196-1b39-64f6f365c1b0/AR_D_annual_report_2020.01.pdf		2021
Video “Twins. The meaning of resemblances and differences for public health”, 1st prize of Amsterdam Public Health Vlogging competition https://youtu.be/S0CKUdGJcR4		2021
Video “The Mystery of Identical Twins and the Epigenetic Signature”, supported by Amsterdam Reproduction & Development https://youtu.be/UyK6hnF7KRA		2022
Video “Science in Shorts”, Nature Awards Competition https://youtu.be/rvWaMzKu_yI		2022

WORDS OF GRATITUDE

Every journey is endowed by people who we encounter, and who open us up to new opportunities and lessons.

The world around me has drastically changed during the past four years of my PhD trajectory in the Netherlands. What has remained is my passion for science and my search for true scholars. I feel enriched by this journey, having had an opportunity to create my own project, and what is now complete, will forever be my personal contribution to science.

I am immensely thankful to my supervisors who were amazing mentors during my PhD projects: Dorret Boomsma, Conor Dolan, Jenny van Dongen, and Toos van Beijsterveldt. It was especially inspiring to observe and learn from their work ethic, dedication to science, a high level of professionalism, and scientific style and diplomacy.

I learned a lot about the world of science from Dorret, who opened many infinitely valuable opportunities for me, for example by encouraging me to participate in professional educational courses and events, acquire teaching experience in her Complex Trait Genetics course, and sharing professional connections. Furthermore, she embraced my creative ideas on translational science and video production. She also helped me in domains beyond academia, for example by helping me deal with myriad challenges during the COVID-19 pandemic and in a time of great uncertainty. Her mentorship will always hold a warm place in my heart. She believed in me...

Rigorous lessons in methodology from Conor formed a solid statistical foundation for my future development. He was a great mentor with just the right amount of patience, and I thank him very much. Together with Jenny, we did a vast amount of work analyzing whole epigenome data. Jenny's inspiration, careful approach, and impeccable academic style together with her balanced approach to life helped me boost my skills and stay on track, despite the challenging nature of the analysis. Thank you, dear Lisa, for bringing more joy in our collaboration with your appearance!

Toos van Beijsterveldt, together with Lannie Ligthart and Gonneke Willemsen provided a wonderful example of how to deal with large complex datasets from the Netherlands Twin Register, which gives us a big leg-up in research. Fiona Hagenbeek made a fantastic work on ACTION data collection that is contributing to epigenetic studies for several generations of researchers, and I am grateful to be among them.

I warmly thank the reading committee that provided their feedback on my PhD thesis so promptly and thoroughly: Elsje van Bergen, Bas Heijmans, Hamdi Mbarek, Tessa Roseboom, Carl Schuengel, and Petra Zwijnenburg.

I would like to mention the people and organizations who supported me in starting the PhD project, and aided in my development in European academia. Thanks to Gian Carlo Di Renzo for introducing me to the Dutch group of twin researchers, and to Carl Schuengel for helping to come to the Netherlands. With deep respect, I thank Gennady Sukhih, who is the director of Kulakov National Center of Perinatology and Obstetrics, to

Victoria Dmitrieva of the Antonio Meneghetti Scientific Foundation, and to Bulat Nureev of the Skolkovo School of Management for supporting the start of my trajectory in Europe.

I would like to extend a thank you to the Amsterdam Reproduction and Development institute, Amsterdam Public Health institute, and the European network staff eXchange for integrAting precision health Care sysTems (ExACT) consortium (Horizon2020), who supported me in several projects that resulted in multiple articles, videos, and fruitful collaborations. And Carla van El, thank you for the collaboration!

During my PhD training at the Vrije Universiteit (VU) of Amsterdam, I learned a lot from wonderful mentors including Michiel Nivard, Meike Bartels, Hamdi Mbarek, Eveline de Zeeuw, Camelia Minica, Elsje van Bergen, Matthijs van der Zee, and Eco de Geus. Jouke-Jan Hottenga and René Pool introduced me to the world of genome-wide association studies and polygenic risk scores, which partially inspired my subsequent choice of a post-doc research project on the genetic architecture of internalizing disorders. And Camelia, thank you for proposing a discussion around the Gattaca movie, a discussion that students in our course found valuable and fun. With Eco de Geus and Martin Gevonden, we discussed many of the valorization aspects of translational research, from which I learned how to make science accessible for everyone, and I was excited to witness the growth of their startup. The help of Natascha Stroo and Michiel Verburgh favored many processes and communications within the VU.

The PhD research environment at the department of Biological Psychology was a fruitful ground for boosting my skills, gaining valuable knowledge, and realizing my creative initiatives. Therefore, I would like to thank Iryna, Camiel, Wonu, Eshim, Margot, Perline, Sofieke, Zenab, Lianne, Susanne, Bodine, Denise, Hill, Floris, Nikki, Yayouk, and Bart.

The freedom to combine the hobbies with research, and the work-life balance modeled by the department gave me new insights on life. For example, the hobbies of several of those in the department included: painting (Toos), guitar (Lannie, Jouke-Jan, Rene), philosophy (Gonneke), athletics (Elsje), korfbal (Nikki), and kungfu (Hill and my wushu school of Shaobo Tang), and these hobbies modeled the wide range of interests that others in my field have.

We were also lucky enough to have the ability to create wonderful videos about our findings on epigenetic signatures in monozygotic twins together with Nikki, Floris, Margot, Denise, Camiel, and science visualization expert Can Lockman (who also generously helped me to design this book). The Epigenetics Journal Club has developed a lot since, while in the beginning it was just the three of us: Jenny, our Master's student Valerie, and me. Observing its current developments and growth is inspiring!

I would also like to thank the Research Master students. It was an honor to teach them during the Complex Trait Genetics course, and several were inspired to continue further collaborations with us. Thank you, Jana, Sophie, Kirsten, Dmitry, Penghao, Emily J., and Yasmin for your contributions to the research projects.

During my stay with Prof. Caroline Relton's group in Bristol, one of the leading epigenetic groups in the world, I gathered substantial experience with epigenetic data analysis.

With Drs. Matthew Suderman and Dorretta Caramaschi, we analyzed data from two large cohorts: the NTR and ALSPAC. Thank you for this dynamic and fruitful part in my research!

I am grateful for our honest discussions around open science, their support, and the research ideas that were developed in my collaborations with Drs. Rosa Mulder and Gennady Roschupkin (Erasmus University), Rens van de Schoot (Utrecht University), Peter Roetman (Leiden University Medical Center), Steve van de Weijer (NSCR), Oxana Naumova (Houston University), and Jeff Beck and Brandon Johnson (Avera Institute for Human Genetics).

I was able to develop ideas on translational projects and valorization through the opportunities and resources provided by the VU. Together with colleagues from the Avera institute, Prof. Erik Ehli and Austin van Asselt, I created an Epigenetic Biomarker Analysis (EBA) project, which was selected for the incubation program at Amsterdam Center of Entrepreneurship. Many thanks to all of the experts who helped me in this perspective, including my encouraging mentors Asya Gyniatullina and Alfred Nijkerk, and successful biotech entrepreneurs Niccola Bonzani, Andrea Candelli, and Thomas Wurdinger for sharing their experiences and ideas through their inspiring talks.

I am immensely happy that the methodological issues of my thesis have already found their place in Qatar, thanks to the initiative of medical doctors and twins Kholoud and Mashael Al Shafai, Qatar Genome researcher Hamdi Mbarek, and Qatar Biobank director Nahla Afifi. It is amazing to see this development, and I believe that it will increase the diversity and generalizability of twin research.

Twin researchers Nick Martin and Nancy Segal were openhearted and empathetic towards me, involving me in very interesting projects, and giving me a chance to write reviews on amazing books on twins. I am also grateful to the experts in early life development: Profs. Tessa Roseboom (the Netherlands), Nils Lambalk (the Netherlands), Judith Hall (Canada), and expert in handedness Chris McManus (UK), who made time to speak with me personally, ultimately fueling the inspiration on many issues described in my thesis. I certainly hope that we will cross roads again someday.

Great professional and journalist Emile Schelvis is an expert in media-management, and guided me through the art of science communication. Together with Emile, we created several successful videos meant to communicate research findings to lay audiences, created a video production team, and organized impressive interviews with scientists, as well as famous twins and their parents. My special thanks go to Gitta de Boer, the mother of Frank and Ronald de Boer, whom I interviewed for a video that won the 1st prize at the Amsterdam Public Health Vlogging Competition, and to Alexander Brink who supported all my creative and scientific ideas, and performed a beautiful piano solo to illustrate the concept of epigenetic signatures to a broad audience.

I would also like to thank Frederik de Vos for his wisdom about art and literature, and for sharing his artistic view of the world with me.

In the background there was of course always my family in Russia, and the Dutch side of the family, who offered their continuous support. They include my aunty Galya and her husband Henk van de Bunt, my cousin Rimma, and friends from Rotterdam, the family of Emile and Michel Schelvis. They introduced me to the Dutch culture and traditions, as well as philosophy and music. Henk, being a Professor of Criminology, offered great insights about science, and was often the first to read my articles together with Galya. I sincerely hold all the wonderful events we shared together in my heart, including their insightful observations about life.

Last but not least, I would also like to extend a heartfelt thanks to all the twins who gave their time to contribute to the advancement of science all over the world, and especially to Frank and Ronald de Boer, Rene and Willy van de Kerkhof, Peter and Sjoerd, Brit and Evi, as well as Lucas, for their help with science communication.

You were all so great.

And the journey will go on.

The great art of living science.