Methodological considerations in epidemiological studies in perinatal medicine

Dominika Modzelewska

Department of Obstetrics and Gynaecology,
Institute of Clinical Sciences
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2021

Cover illustration by JBD

Methodological considerations in epidemiological studies in perinatal medicine

dominika.modzelewska@gu.se

ISBN 978-91-8009-326-2 (PRINT)
ISBN 978-91-8009-327-9 (PDF)

Printed in Gothenburg, Sweden 2021
Printed by Stema Specialtryck AB, Borås, Sweden, 2021

'If I speak in the tongues of men and of angels,
but have not love,
I am only a ringing gong or a clanging cymbal.'

I Corinthians 13:2

Methodological considerations in epidemiological studies in perinatal medicine

Dominika Modzelewska

Department of Obstetrics and Gynaecology, Institute of Clinical Sciences
Sahlgrenska Academy, University of Gothenburg
Gothenburg, Sweden

ABSTRACT

This thesis is based on two projects, each consisting of two papers. In these two projects we investigated factors affecting the variability in gestational duration, and associations between maternal diet and neonatal outcomes.

In the first project we investigated factors affecting the variability in gestational duration with a particular interest in the genetic contributions. The contribution of genetic factors to trait's variability is estimated by heritability analysis. For gestational duration, heritability estimates are up to 36%. However, genetic studies have identified a limited number of genetic variants. In the first project we aimed to assess whether an essential assumption underlying heritability analysis, constant environmental conditions, holds. In Paper I, we observed a substantial variability in the correlation estimate in gestational age at delivery between relatives with respect to their year of birth or age gap. This variability suggests the existence of temporally changing environmental factors. In Paper II, we showed that obstetrical practices and data handling contribute to changes in gestational duration over time. Two studies in this project suggest that the assumption of constant environmental conditions might not hold. This might explain the dissonance between the conclusions of heritability and genetic studies.

The second project was an extension of previous studies performed in our research group. In these studies, we detected associations between maternal

caffeine or selenium intakes and small for gestational age (SGA), and SGA and neonatal outcomes. Therefore, we hypothesized that the selected maternal food components and neonatal outcomes would be associated. In both studies (Papers III and IV), we did not detect associations between maternal caffeine or selenium intake and neonatal outcomes. The lack of associations may be the result of SGA babies having different neonatal outcomes based on the underlying SGA cause. One type of SGA might be related to maternal caffeine/selenium intake and not associated with neonatal outcomes; the other types might be caused by more aggressive factors leading to worse performance in the neonatal period. The results raise caution for using SGA as a representation for neonatal outcomes.

Keywords: gestational duration, preterm delivery, familial aggregation, heritability, variability, birthweight, small for gestational age, neonatal outcomes

ISBN 978-91-8009-326-2 (PRINT)
ISBN 978-91-8009-327-9 (PDF)

Sammanfattning på svenska

Denna avhandling är baserad på två projekt, som vardera består av två artiklar. I dessa två projekt undersökte vi dels faktorer som påverkar variationen i graviditetslängd, dels samband mellan moderns kost och neonatala utfall.

I det första projektet undersökte vi faktorer som på verkade variationen i graviditetslängd. Vi var särskilt intresserade av genetiska faktorers påverkan på graviditetslängdsvariationen. För graviditetslängd förklarar ärftlighet upp till 36% av variationen. Genetiska studier har dock identifierat ett begränsat antal genetiska varianter. I det första projektet syftade vi till att utröna om ett grundläggande antagande bakom släktskapsanalys och konstanta miljöförhållanden är riktiga. I artikel I observerade vi en betydande variation i korrelationsuppskattningen av graviditetslängd vid förlossningen mellan släktingar, med avseende på deras födelseår eller åldersspann mellan graviditeterna i registret. Denna variation kan tyda på att det finns temporärt förändrade miljöfaktorer. I artikel II visade vi att obstetrisk praxis och datahantering bidrar till förändringar i graviditetslängden över tiden. Två studier i detta projekt antyder att antagandet om konstanta miljöförhållanden kanske inte håller. Detta kan förklara dissonansen mellan slutsatserna gjorda om ärftlighet och genetiska studier.

Det andra projektet var en utökning av tidigare studier utförda i vår forskargrupp. I dessa studier upptäckte vi samband mellan koffein- eller selenintag från mödrarna och om barnet var lätt för tiden (SGA) samt SGA- och neonatala utfall. Därför antog vi att utvalda komponenter i mammornas kost och nyföddas hälsa var associerade. I båda studierna (Artikel III och IV) upptäckte vi inte samband mellan mödrars koffein- eller selenintag och neonatala utfall. Bristen på associationer kan vara resultatet av att SGA-barn har olika neonatala utfall baserat på den underliggande SGA-orsaken. En typ av SGA kan vara relaterat till mödrars koffein / selenintag och inte associerat med neonatala utfall; de andra typerna kan orsakas av mer sjukdomsframkallande faktorer som leder till sämre barnutfall under nyföddhetsperioden. Resultaten visar på att man skall använda SGA med försiktighet som riskfaktor för dåligt neonatalt utfall.

Nyckelord: graviditetslängd, för tidig födsel, familjär aggregering, ärftlighet, variation, födelsevikt, liten för graviditetsålder, neonatalt utfall

List of papers

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. Modzelewska D, Sole-Navais P, Zhang G, Muglia LJ, Nilsson S, Jacobsson B. *Importance of the environment for gestational duration variability and correlation between relatives - results from the Medical Swedish Birth Registry, 1973-2012.* PLoS One. 2020 Jul 24;15(7):e0236494.

II. Modzelewska D, Sole-Navais P, Sandstrom A, Zhang G, Muglia LJ, Flatley C, Nilsson S, Jacobsson B. *Changes in data management contribute to temporal variation in gestational duration distribution in the Swedish Medical Birth Registry.* PLoS One. 2020 Nov 6;15(11):e0241911.

III. Modzelewska D, Bellocco R, Elfvin A, Brantsæter AL, Meltzer HM, Jacobsson B, Sengpiel V. *Caffeine exposure during pregnancy, small for gestational age birth and neonatal outcome - results from the Norwegian Mother and Child Cohort Study.* BMC Pregnancy Childbirth. 2019 Feb 26;19(1):80.

IV. Modzelewska D, Solé-Navais P, Brantsæter AL, Flatley C, Elfvin A, Meltzer HM, Sengpiel V, Barman M*, Jacobsson B*. *Maternal Dietary Selenium Intake during Pregnancy and Neonatal Outcomes in the Norwegian Mother, Father, and Child Cohort Study.* Nutrients. Multidisciplinary Digital Publishing Institute; 2021 Apr;13(4):1239

---

* *Shared last authorship*

Content

ABBREVIATIONS

Abbreviations

BMI	Body mass index
GWA	Genome-wide association
h^2	Heritability estimate
LBW	Low birthweight
LMP	Last menstrual period
MBR	Swedish Medical Birth Registry
MBRN	Medical Birth Registry of Norway
MoBa	Norwegian Mother, Father, and Child Cohort Study
PL	Preterm labor
PTD	Preterm delivery
PROM	Prelabor rupture of membranes

01 RESEARCH PROJECTS

1 Research projects

1.1 Projects included

1.1.1 Research fields

This thesis includes two research projects, each consisting of two papers, in the fields of genetics (Project 1) and nutrition (Project 2) in pregnancy. Within those fields, we focused on the variability of pregnancy duration and the effect of maternal nutrition on newborn's health. Pregnancy duration and maternal nutrition represent significant components of pregnancy upon which fetal development depends on. To develop properly, a fetus requires both time and an appropriate environment. It is estimated that a fetus needs around 40 weeks to develop properly. These weeks must be accompanied by a proper intrauterine environment which is achieved, among others, by an adequate maternal supply of nutrients.

1.1.2 Project 1

In the first project, we aimed to contribute to the understanding of the nature of pregnancy duration by exploring familial aggregation. We attempted to understand the genetic basis by which relatives (for example, sisters) have pregnancies with similar duration. In other words, we were interested in understanding the dependence of pregnancy duration on inherited genes. Heritability analyses are used to assess the extent to which variability in the phenotype depends on genes transmitted from parents to offspring. The heritability estimate is generally interpreted in terms of the relative importance of genetics for a given phenotype. One of the core assumptions that must hold for heritability analysis to be valid is a constant environment. This assumption was a starting point for the first two articles included in this thesis.

1.1.3 Project 2

In the second project, we concentrated on the importance of maternal nutrition for fetal development reflected in the baby's health in his first 28 days of life. In particular, we focused on two common food components: caffeine and selenium. The project was born as an extension of previously conducted studies in our research group[1–3]. Our colleagues observed that maternal

consumption of both caffeine and selenium is related to baby's birthweight. Baby's birthweight is often used as a proxy for other neonatal health complications. Babies with a birthweight smaller than expected have higher chances to die or have neonatal morbidities. We presumed that studying the relationship between the above-mentioned food components and neonatal outcomes would offer insights into the nature of birthweight. That expectation became a departure point for the next two studies included in this thesis.

1.2 What is this thesis about?

The aspects covered in this thesis are broad and might be tackled in multiple directions. In both projects, the overall goal was to get insights into the nature and etiology of perinatal outcomes: pregnancy duration and birthweight. I narrowed the discussion to the puzzling observations in the fields of genetics of pregnancy duration and birthweight, known under the terms of 'missing heritability'[4] and 'birthweight paradox.'[5]

02 STUDY FILEDS

2 Study fields

2.1 Introduction

The fields of human genetics and nutrition may provide explanations on the factors determining the variability in phenotypes observed in human life. Individual genetic makeup forms the basis of their lifetime health traits. In the era of genomics, among other -omics, we took a step back to take a picture of the genetics of pregnancy duration from a larger angle. While genetics define a large fraction of the variability of a multitude of traits (height, body mass index (BMI), or several psychiatric disorders), the contribution of genetics to pregnancy traits is generally smaller. For such traits, the maternal environment plays a crucial role. For the offspring, specific exposures at stages of development when the biological plasticity is the highest might be particularly important for lifelong health. Nutrition may bring permanent changes to the metabolism or physiology of the organism when tissues and organs develop[6]. Genes transmitted to the fetus from parents and the intrauterine environment, affected by maternal diet, are involved in determining an individual's short- and long-term health condition.

2.2 Quantitative genetics

2.2.1 Gene – heredity unit

Before genes were recognized as heredity units, it was believed that similarity between parents and offspring is the result of parental bloods being mixed in the offspring[7]. The belief of heredity being strictly associated with “blood” was abolished after the observations of the founder of modern genetics, Gregor Mendel.

Gregor Mendel (1822 – 1884) was a mathematician and biologist and is known for his study done in his monastery’s garden. In his experiment, Mendel crossed peas with multiple different characteristics (seed’s and pod’s colors and shapes, plant height, flower color, and location). As a result, he observed a pattern in which the parental peas’ characteristics were co-occurring in the next generation. Clear and repetitive distribution of characteristics gave birth

to the concept of heredity units, which, many years after Mendel's death (in 1909), Wilhelm Johansson called genes[8].

Mendel observed that the variation in inherited characteristics depends on the presence of genes' alternative forms (called alleles, dominant or recessive) in the parental cohort. He concluded that the organism inherits two alleles, one from each parent. As it was recognized by himself and biologists of the time, the results were not seen as generally applicable, but only to some species or traits[7]. Mendel's laws of inheritance apply to traits that are determined by a single gene (at a single position at the genome, called locus), and have easily distinguishable forms (i.e., qualitative traits).

2.2.2 Mendelian traits

Mendelian disorders might be recognized by occurring in families. Examples of Mendelian disorders are color blindness, cystic fibrosis, or hemophilia[9]. All those disorders depend on the expression of a single gene, meaning that the person with a mutation reducing the function of that gene also has a given phenotypic value for the disorder. The discipline responsible for studying the inheritance of Mendelian traits is called Mendelian genetics.

2.2.3 Polygenic (complex) traits

Most of the traits are not defined solely by one gene but are the product of cumulative or interacting actions of multiple genes and non-genetic factors. Traits defined by multiple genetic and environmental factors are called quantitative or complex. Quantitative traits do not follow Mendelian ratios but might vary among individuals, producing continuous distributions of phenotypes. The discipline studying the inheritance of quantitative traits is called quantitative genetics[10]. As the extension of Mendelian genetics, this discipline relies on Mendel's proposition of similarity among relatives being determined by genetic transmission. At the time of Mendelian and quantitative genetics establishment, genes were recognized as the sole unit accounting for the phenotypic similarity between relatives[7]. Today, more factors contributing to the phenotypic similarity among relatives are acknowledged, such as microbiome[3] or life-style[4].

2.2.4 Quantitative genetics vs. Mendelian genetics

While both Mendelian and quantitative genetics aim to contribute to the understanding of phenotypic inheritance, they differ in the subject of study (Table 1). In Mendelian genetics, a single gene being cause for the phenotypic value creates individuals carrying the same genetic variant having the same characteristic; therefore, phenotypic similarity is the subject of study. In quantitative genetics, multifactorial and complex etiology of phenotypes results in population variability; thus, phenotypic variability became the subject of study. Given the difference in the subject of study, the questions to be answered in those fields are different. In Mendelian genetics, the aim is, once the causal gene is identified, to determine the inheritance pattern of the single genetic variant, while quantitative genetics assesses the size of the relationship between genetic variability among individuals and phenotypic variability. Different scientific questions require different study units, being families and populations in Mendelian and quantitative genetics, respectively.

	Mendelian genetics	Quantitative genetics
Factors determining phenotype	Single gene	Multiple genes, non-genetic factors
Factors underlying phenotypic similarity	Genes	Genes, shared environment
Scientific question	What are the causal gene and the inheritance pattern?	What is the contribution of genetic variability among individuals to phenotypic variability?
Subject of study	Phenotypic similarity	Phenotypic variability
Unit of study	Pedigree	Population, pedigree
Examples	Sickle cell anemia, color blindness	Height, weight

Table 1. Differences and similarities between Mendelian and quantitative genetics.

2.2.5 Heritability analysis

Quantitative genetics study the genetic profile of a population in order to recognize the genetic basis of the phenotypic variability[10]. One of the main measures in quantitative genetics is the heritability estimate. Heritability is defined as a ratio of additive variance to phenotypic variance and is interpreted as the extent to which transmitted parental genes determine the phenotype of their children[11]. Therefore, heritability estimate summarizes the phenotype's dependence on genetics[12] and is identified as a valuable tool in disease risk prediction[11].

The heritability estimate is retrieved from the phenotypic similarity between relatives. Relatives are similar to each other due to shared factors, genetic and environmental. Different types of relatives share different numbers of those factors to a different degree, affecting the extent to which relatives resemble. The heritability estimate is obtained based on the assumed composition of the factors accounting for the phenotypic similarity (Figure 1). The composition is valid under the assumption of constant environmental conditions, population being in Hardy-Weinberg equilibrium, and additive effects of genes[10].

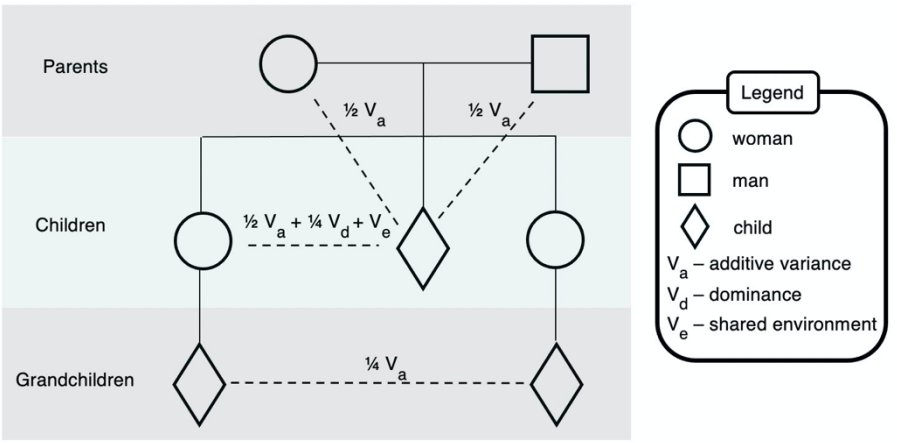


Figure 1. Simplified pedigree with marked composition of factors underlying the phenotypic similarity between relatives.

The heritability estimate, as a relative value, ranges between 0 and 100%. It has been accepted that high heritability estimates are an indicator of the phenotype of an individual being a good predictor of the genotype in the

current population[12]. Therefore, a low heritability estimate does not eliminate the importance of genes but indicates that a small portion of phenotypic variance is caused by variation in genotypes[12]. Such understanding leads to searching populations with high heritability estimates as it gives promise for genome-wide association (GWA) analysis to detect genetic variants when performed in those populations.

It has been acknowledged that heritability estimates vary across studies. That variation is accepted on the basis of the property of heritability being a population parameter[12]. Heritability estimates reflect the genetic and environmental variance of a specific population at the specific time. Therefore, any change in the genotypic profile or environmental conditions is expected to be reflected in the heritability estimate.

2.2.6 Missing heritability

Heritability analysis has provided large estimates for multiple traits, such as 80-90% for height[4], 90% for autism[13], or 80% for schizophrenia[14]. Large estimates gave promises for successful genetic variant detection. However, GWA findings do not account for the totality of the heritability estimates. GWA studies identified numerous genetic variants but with tiny individual or cumulative effects; therefore, of no prediction usage[11]. That dissonance between the findings and conclusions of heritability estimates and genetic association findings had puzzled scientists since 2008 when the conflict was officially termed as *missing heritability*[4].

Possible reasons for the occurrence of missing heritability are related to the limitations of GWA studies[4]. The rationale for performing GWA analyses is based on the hypothesis on the etiology of common diseases. As for a single gene disorder, one single mutation is enough for a person to develop the disease; for a complex disease, it was hypothesized that multiple genetic variants underlie the traits. However, GWA analyses brought less information than expected, consequently questioning the legitimacy of the hypothesis[4]. On the other hand, it is possible that small effects of genes are undetectable in the currently available DNA datasets. Furthermore, GWA studies are not accounting for the possible interaction between genes resulting in a lower variance explained than could be expected based on heritability estimates[4]. Finally, heritability estimates might be biased due to not accounting properly for the environment[4].

2.3 Nutrition

2.3.1 Perinatal nutrition

The understanding of the importance of maternal nutrition during pregnancy has been changing over the years. In the 1950s, malnutrition was considered to affect only the mother and not the fetus. As a "parasite," the fetus was assumed to be protected by the mother from any nutritional shortages[15]. Required nutrients were anticipated to be supplied from maternal body composition. Such a perspective was based on two major studies of that time that reported a lack of effect of maternal supplementation during pregnancy on neonatal outcomes[16–19]. Maternal nutrition during pregnancy gained greater consideration after an increased understanding of fetal development[20]. It was theorized that there are critical time windows in fetal development that are sensitive to external factors.

In a natural experiment during the Great Famine, a highly unbalanced diet in the last months of pregnancy was observed to be harmful for the fetus. Records collected from women who during their third trimester of pregnancy experienced hunger showed lower birthweight among their offspring in comparison to mothers who had not been under the famine in the earlier pregnancy periods[21]. Later studies reported the link between birthweight and health status in adult life. Those who had low birthweight had a higher risk of developing coronary artery disease, diabetes, or metabolic syndrome later in their life[22–24].

The effect of maternal intake on fetal development has been studied in animal models. In rats, a reduction in caloric intake by 30% resulted in a substantially lower birthweight of the offspring. Underfed *in utero* offspring had a higher tendency to gain fat faster in their adulthood in comparison to rats whose mothers were fed with a normal diet during pregnancy[25]. Iron deficiency before and during pregnancy among rats was related to altered cardiac development and increased blood pressure in the offspring[26,27]. Increased blood pressure was also observed in the offspring of rats fed with a high-fat diet during pregnancy[28]. High consumption of protein in rats during pregnancy was associated with diminished energy expenditure and obesity[29].

Maternal nutrition is considered to introduce permanent changes to the metabolism and physiology of the individual by interfering with tissue and organ development[6]. The process of tissue and organ formation starts from a small group of cells in the embryo. Those cells undergo growth and division (proliferation process) and specialize into specific cell types (differentiation process). Disruption of those processes due to unfavorable intrauterine conditions might lead to a reduced organ size (if affected during proliferation stage) or altered cell function (if affected during the differentiation stage)[6].

The effect of maternal nutrition on the fetus is dependent on the efficiency of the placenta. The placenta is the organ responsible for the nutrient and oxygen transport between mother and fetus. Proper nutrient flow is dependent upon the placenta's invasion of the uterus tissues (placentation), the size, and vascular system development[30]. The blood flow is controlled by the hormones produced in the placenta, and its secretion might be, in turn, reduced by maternal undernutrition[19]. Weak placentation or uteroplacental blood flow may result in limited nutrient transport and, consequently, restricted fetal growth and development.

2.3.2 Selected food components

In the papers included in this thesis, we studied maternal intake of caffeine and selenium.

2.3.2.1 Caffeine

Caffeine is largely available and consumed worldwide. Coffee and tea are the main sources, but caffeine is also found in energy or soft drinks and chocolate. On average, a cup (250 ml) of brewed coffee contains 94 mg of caffeine, tea 50 mg, and chocolate 7mg[31]. Caffeine content varies in energy drinks and might be up to 175 mg per 250 ml.

Caffeine is commonly acknowledged to affect mood, stimulate perception, cognition, or behavior[32]. It takes between 30 and 60 minutes for caffeine to absorb fully and arrive at its peak level in the blood[33,34]. It is commonly used by athletes before sport's competitions to increase their performance[35]. Caffeine has been found to be more effective among trained individuals than untrained[36]. Complex supplements with caffeine are improving strength, endurance, and lean body mass[36]. The dosage of 3 to 9

mg/kg of caffeine, translated to 210-630 mg/day for a 70-kg person[37], is considered enough to observe improvements in physical performance.

Caffeine intake below 400 mg/day is considered moderate and not having adverse effects, and 10g has been estimated to be lethal[38]. Epidemiological studies showed that moderate caffeine consumption might prevent Alzheimer's or Parkinson's diseases, while evidence gathered from animal studies suggest that caffeine might play a therapeutic role among already diagnosed patients[39]. Beneficial effects of caffeine consumption were also observed for type-2 diabetes, several types of cancer (e.g., oral, liver), or chronic liver diseases[40].

Adverse effects of caffeine have been reported in studies on pregnancy outcomes and might be attributed to decreased caffeine clearance during pregnancy[41]. The variability in the effect of caffeine on the reproductive system suggests a variability in an individual's sensitivity toward caffeine[12]. A daily consumption of 300 mg of caffeine was related to a higher risk of pregnancy failure and low dosage between 100 and 200 mg with miscarriage, fetal growth restriction, or low birthweight[38,42–44]. Individuals exposed to caffeine during intrauterine life had a higher risk for impaired cognitive development and obesity during childhood[45–47].

2.3.2.2 Selenium

Selenium is an element present in the soil and groundwater. The availability of selenium varies over geographical areas and might be affected by industrial or agricultural practices (industrial processes of extracting metals or agricultural water systems applications)[48,49]. The highest soil content of selenium has been found in some parts of the US, China, or Ireland, and the lowest in Finland, Serbia, or Congo[50,51].

Variability in the selenium availability in the environment leads to variability in the selenium content in foods. The main sources of selenium for a human are wheat and other grains and meat. Primarily, selenium is absorbed in the small intestine and in the blood is transported by binding to proteins. It is found in multiple tissues, is excreted in urine, feces and expired with air.

Recommended dietary allowance (RDA) varies between countries. In Nordic countries, RDA is 50 μg/day for females and 60 μg/day for males[52]. During pregnancy and lactation period, consumption of 60 μg/day of selenium

is recommended[52]. RDA in countries such as Germany, UK, Australia, or the USA is similar to Nordic countries[53–56]. Lower intakes are also advised; for example, in Japan, 25 μg/day for women and 30 μg/day for men[57]. However, the tolerable upper intake level of selenium intake is similar among all mentioned countries and varies between 300 and 450 μg/day[57].

The overdose of selenium is rare and might happen in countries with rich soil content or due to over-supplementation[56,58]. The lowest selenium intake level at which first adverse effects are occurring is 910 μg/day[59]. Overdosing relates to nausea, vomits, diarrhea, tachycardia[60]. Selenium deficiency is much more common, being recognized among 500 million to 1 billion people globally. Selenium deficiency relates to issues with heart (Keshan diseases), joints (Kaschin-Beck disease), and mental health (myxedematous cretinism)[61].

Selenium blood concentration is lower during pregnancy. On average, selenium concentration is at 35 μg/L and 59 μg/L among pregnant and non-pregnant women, respectively[62,63]. Selenium deficiency has been associated with impaired development of the nervous system[64].

03 PERINATAL OUTCOMES

3 Perinatal outcomes

3.1 Introduction

The success of pregnancy is assessed based on different measures related to a newborn's biological maturity and developmental degree. Pregnancy development can be assessed at different stages; during pregnancy, around delivery, or during the newborn's first 28 days of life. Respectively, measures collected at those stages are called prenatal, perinatal, and neonatal outcomes.

The outcomes in three pregnancy time periods (prenatal, perinatal, and neonatal) are closely connected. During the prenatal period, the effects of maternal health condition and fetal development might be observed in perinatal outcomes. Subsequently, perinatal outcomes might be reflected in the neonatal measures. All those outcomes allow us to predict an individual's health from his first days after birth.

Which of the three pregnancy-related periods is the most important in the perspective of the baby's health and development later in life? The answer is not simple, and multiple aspects must be taken into account. The same outcome's value might be a result of different processes and might shape an individual's later development in different ways. Additionally, an individual's development depends on the mechanism of catching-up for adverse outcomes. That complex and multi-layered interdependence suggests that the best strategy should focus on the earliest stages in which a more determinant intervention can have a bigger impact. Hence, the study of perinatal and neonatal outcomes might give insights on the importance of factors that should be prioritized during pregnancy for improving human health.

3.2 Gestational duration

3.2.1 Definition

Gestational duration refers to the period of time a woman is pregnant. It is the interval between conception and delivery. In the literature, another very common term is "gestational age at birth," in short, "gestational age." Those

terms are used interchangeably. However, given that there is no strict consensus on the terminology in this thesis, those terms are differentiated.

In this thesis and throughout the papers, "gestational duration" is used to express duration as a characteristic of a woman or pregnancy (thus, as a biological phenomenon). It describes a time-interval rather than a single time point (Figure 2). Oppositely, "gestational age at birth" is used to express gestational duration as a characteristic of an individual. It refers to a specific time point (measured on the scale of gestation duration) when the specific event (birth) occurs (Figure 2). In this thesis, "gestational age" is used to refer to a specific time point during pregnancy.

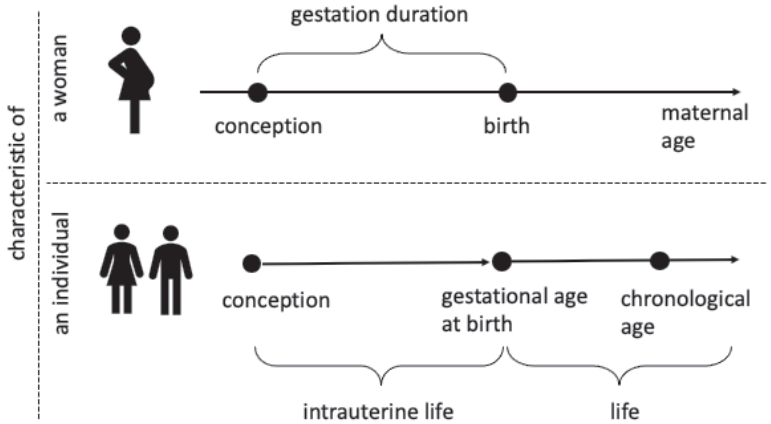


Figure 2. Visualization of the concepts: gestational duration as the characteristic of a woman, gestational age at birth as a characteristic of an individual.

3.2.2 Units

To express gestational duration, the most convenient unit is a day since more compact units result in information loss. Therefore, gestational duration reported in weeks is usually accompanied by the number of days (between 0 and 6). The number of days is reported after the number of weeks followed by the plus sign. For example, the average gestational duration is 279 days, which is reported as 39+6 weeks. Gestational age is also often referred to in trimesters (first trimester, up to 13+6 weeks; second, up to 22+0 weeks, and third, at the end of pregnancy)[65]. However, such specification is more often used in the clinical settings rather than in the research. It is used to specify the best time window for given obstetrics measurements. For example, first or second

trimesters are considered as the best windows for a due date (i.e., delivery day) estimation[66].

3.2.3 Estimation methods

There are two methods available to estimate gestational duration; based on last menstrual period (LMP) date or ultrasound measurements. Both methods are based on the assumption of the length of normal pregnancy, which nowadays is assumed to be 280 days. Gestational duration is calculated by adding to the normal pregnancy the difference (in days) between birth date and expected delivery day (due date), *280 + (birthdate – due date)*[67].

The first step in estimating gestational duration is the determination of the due date. The due date is one of the most important information established during the first antenatal visits of a pregnant woman. This information is fundamental for the appropriate arrangement of obstetric care. The knowledge about a due-date or fetal gestational age allows to properly schedule antepartum tests, assess the fetal growth, and plan proper perinatal or neonatal interventions in case of preterm- (PTD) or post-term deliveries.

Traditionally, gestational age has been estimated from the LMP date. Regular menstruation of 28 days with the ovulation occurring at 14th days is assumed in this method[68]. Estimation based on LMP date relates to the underestimation of gestational age and, therefore overestimation of gestational duration[69]. Also, it has been shown that there is a specific pattern for a woman to report LMP every 5th day of a month. Once the LMP date is known, gestational age at the antenatal check-up and the due date are estimated. In Sweden, before the introduction of electronic journal systems, paper calendar plates ('snurra') were commonly used for the estimation of the due date. Subject to recall and measurement bias, LMP-based estimates have been slowly replaced by estimates based on ultrasound biometry measurements. Initially, due to ultrasound-based estimates being subjected to the variation in fetal growth, LMP-based estimates were preferred with an adjustment for ultrasound when pregnancy duration estimates from LMP were unreliable[70]. Later, ultrasound-based estimates became a standard.

3.2.4 Population variability

The plausible range for durations of pregnancies resulting in a live baby is between 154 and 308 days (22+0 and 44+0 weeks). However, that range might differ between populations. In Sweden, for example, gestational duration was distributed differently in early- and later years (Figure 3). In later years, more common was to deliver at earlier gestational ages than later.

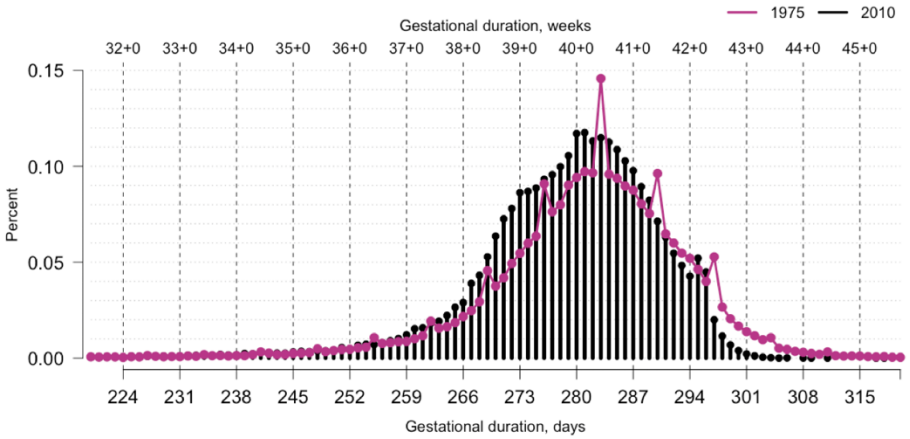


Figure 3. Gestational duration distribution, Swedish Medical Birth Registry. The figure shows the percentage of specific gestational durations from the years 1975 (black line) and 2010 (purple line).

3.2.5 Importance of gestational duration

Gestational duration is considered as the best proxy for fetal maturation and development. Developmental maturity achieved during pregnancy determines the adaptation ability of the newborn to extrauterine life and the neonate's health. From that perspective, the optimal gestational age to be born is between 37 and 42 weeks of gestation (term deliveries). Babies born before 37+0 (preterm) or after 42+0 (post-term) weeks of gestation are more likely to require medical assistance in their first days of life. The relationship between gestational duration and neonatal performance led to the classification of delivery timing into three main categories; preterm, term, and post-term (Figure 4)[68,71].

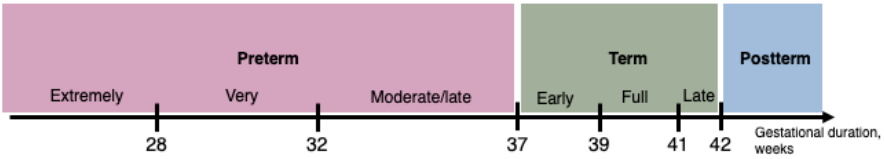


Figure 4. Categorization of gestational duration into preterm, term, and post-term deliveries.

Preterm deliveries: Among preterm-born babies, the risk of death or the type and severity of health complications is dependent on gestational age at birth, with shorter gestational duration leading to more serious health outcomes. With today's medical advancements, babies born as early as 22 weeks of gestation might survive and mature postnatally. However, the majority of deaths (75-80%) in the neonatal period (28 days after birth) are among babies born < 37 weeks of gestation[72–74], and around half of these are among infants born < 32 weeks of gestation[75]. However, even when they survive, babies born before 37 weeks of gestation are at higher risk to suffer from short- and life-term health consequences. Preterm born individuals that survive have a higher risk for neurodevelopmental impairment (from reduced cognitive abilities and behavioral problems to severe mental retardation, cerebral palsy, and epilepsy)[76–78], respiratory problems (such as infections, asthma, bronchopulmonary dysplasia, other chronic lung diseases)[79–82], and visual abnormalities[83,84].

Post-term deliveries: Besides birth complications (umbilical cord complications, shoulder dystocia, peripheral nerve damage, or traumatic injuries), children born after 42nd weeks of gestation have a higher risk of breathing problems (asphyxia, meconium aspiration syndrome) or infections (pneumonia, sepsis)[85,86].

3.2.6 Can gestational duration be controlled medically?

With today's understanding of pregnancy, the duration can be controlled partially. It is easier to invoke delivery rather than to maintain a pregnancy that is prone to spontaneous early delivery. The mechanisms underlying pregnancy maintenance and factors triggering the parturition process are too little understood to build effective preventive methods. Progesterone is a known hormone responsible for maintaining normal pregnancy. The administration of the antagonists of progesterone increases the risk of preterm delivery. However, the administration of progesterone among women is not efficient enough to prevent PTD and decrease its rates[87,88].

Different methods are available to start labor artificially (mechanical/non-pharmacological, pharmacological, or complementary), and when induction of labor is not viable, a cesarean section might be performed. Therefore, past-due date pregnancies might be medically controlled, and post-terms are induced. Until now, there is no consensus on the optimal timing, and more research is required in that area. In Sweden, around 15 to 20% of women have pregnancies lasting 41+0 weeks, and labor is managed at week 42+0[89].

3.2.7 Possible etiologies of preterm deliveries

Preterm delivery is considered as a symptom representing different conditions with different etiologies[90]. Different etiologies might explain the variability in gestational ages at which delivery takes place, different clinical presentation[91], and different pregnancy or neonatal outcomes. Based on the clinical features, preterm deliveries are divided into those without any prior signs of upcoming labor (preterm labor, PL), or with prelabor rupture of membranes (PROM), or those which were started by medical action (iatrogenic).

Preterm iatrogenic deliveries are performed due to pregnancy complications threatening maternal life or fetal development. The main conditions are preeclampsia and severe intrauterine growth restriction, and others include polyhydramnios, oligohydramnios, abruption of the placenta, placenta previas, multiple births[92]. PROM has been associated mainly with maternal infections, while preterm labor with infection, inflammation, stress, hemorrhage, uterine overdistension, or short cervix[93].

Multiple risk factors have been associated with spontaneous deliveries (PL and PROM). Spontaneous deliveries are more frequent among women with extremes of maternal age (below 16 and above 35 years old) or BMI, primiparity,[92] having an imbalanced diet, low socioeconomic status,[91,94–98] smoking during pregnancy, abusing addictive substances (alcohol, drugs), having stress or anxiety during pregnancy,[99] short intervals between pregnancies,[100–102] or preterm delivery history[103], with systemic infections (e.g., pneumonia[104], malaria[105], genital tract infection, intrauterine infection, inflammation[106], with medical disorders (asthma, diabetes, cardiovascular diseases, nephritis, thyroid disease, or chronic hypertension)[103].

While the list of risk factors is long, early prediction of preterm delivery is challenging. More than half of PTD happens in low-risk pregnancies[107,108]. Early delivery signs such as uterine contractions lead to labor in half of women

having them. In another half, without any interference by the obstetrician, contractions stop. Until now, previous preterm delivery in multipara women has been best considered as the most predictive factor. Woman whose previous child was born with PTD has up to 6 times higher risk to deliver the next child prematurely.

3.2.8 Familial aggregation of gestational duration

Familial aggregation, also called familial clustering, of a disease, is the repetitive occurrence of an event within families. Familial aggregation is represented by the risk of the diseases given family history or by the degree of phenotypic similarity.

The observation of a woman having a consistent duration across her pregnancies has been repeatedly presented in the literature. A woman's own history of PTD is currently the best predictor of PTD. Repeating occurrence of PTD within a woman indicates dependence of gestational duration on maternal characteristics or shared environmental conditions in her family. In the early years of the research on familial aggregation of gestational duration, it was suggested that familial aggregation is rather due to factors shared by relatives within one generation and not factors shared between generations[109]. Such conclusion was based on the studies reporting associations between gestational ages at birth of siblings but not among intergenerational pairs[109,110].

Intergenerational relationships, also known as transgenerational, is the relationship between the characteristics of parents and their offspring. Intergenerational association points toward the genetic background of a trait. For gestational age at birth, the first significant mother-child association was reported in 1997 by Porter et al.[111] However, later reports provided mixed results[21–28]. It was suggested that bigger pedigrees should be studied in order to spot the possible genetic inheritance mode[112].

Pedigree studies suggested that, if genetic inheritance existed at all, it ran only in the maternal line. Increased risk for a woman to deliver preterm was observed if her mother or sister experienced preterm delivery, but no risk if PTD occurred in the family of her father or partner[113].

3.3 Birthweight

3.3.1 Classification

Birthweight is the weight registered within hours of an infant's birth[114]. It is an easy and quick measurement but otherwise powerful tool[115]. It reflects a baby's health condition and is used to predict survival chances during the first month or later development[115]. The reverse relationship between baby's birthweight and mortality/morbidity (lower birthweight higher risk) led to subgrouping babies into those with low birthweight (< 2.5kg, LBW), very low birthweight (< 1.5kg), and extremely low birthweight (< 1kg).

On average, birthweight among live singletons born at term might vary between around 3kg and 4kg[116]. In Sweden in years 2010-2012, the average birthweight was around 3.5kg, and 83% weighted between 3kg and 4.5kg. Only 3% of children were LBW (Figure 5).

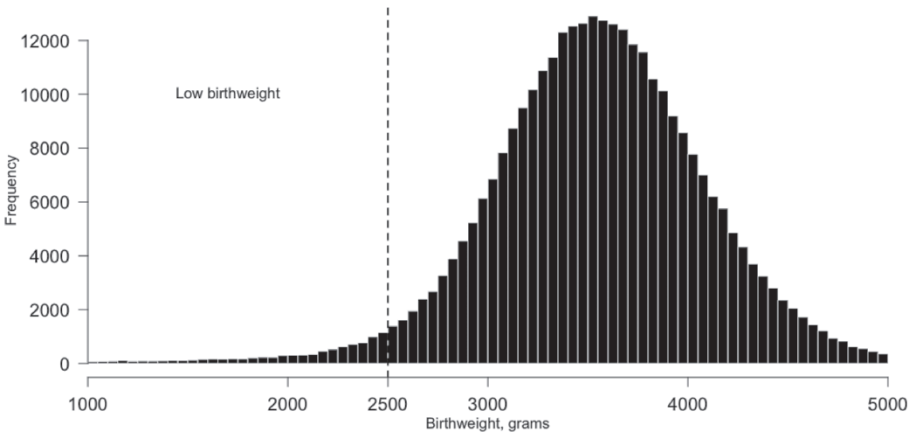


Figure 5. Birthweight histogram, Swedish Medical Birth Registry 2010-2012, number of births: 318 373.

3.3.2 Fetal growth

Baby's birthweight is reflecting its growth in the maternal uterus. Fetal growth rate might be restricted due to some unfavorable intrauterine conditions[117]. To a large extent, fetal growth is dependent upon adequate placenta functioning. Improper placenta's attachment to the uterus or poorly developed vascular system increases the risk of limiting fetal growth[118].

Birthweight considered alone might provide incorrect conclusions about fetal growth. Not all children with low birthweight had disrupted growth in utero. A fetus could have a normal growth rate but born prematurely will weigh less than 2.5kg. On average, children born at 244 days of pregnancy weigh around 2.5kg (Figure 6). In such a case, low birthweight is a result of prematurity and not pathological processes affecting growth. Birthweight is gestational duration-dependent; therefore, gestational duration-adjusted birthweight is more informative.

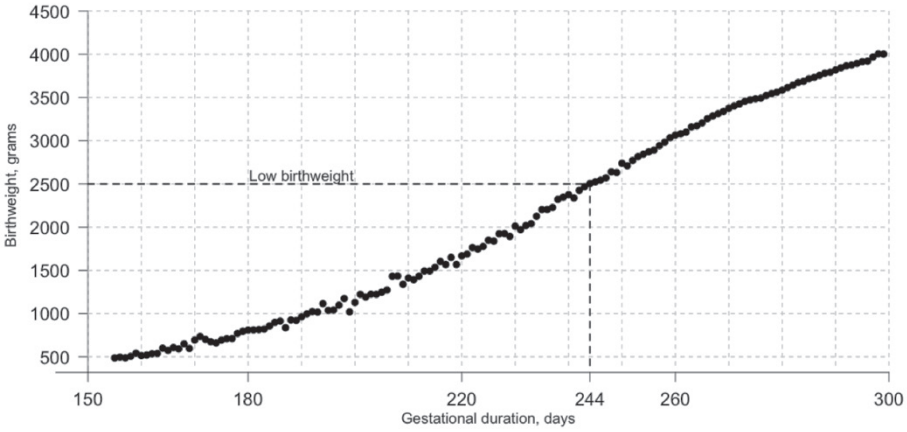


Figure 6. Average birthweight for gestational duration, Swedish Medical Birth Registry 2010-2012, number of births: 318 373.

3.3.3 Small for gestational age

For each gestational age and sex, there is physiological variability in the normal fetal size, which is dependent on fetal growth potential. Fetal growth potential is genetically determined, but growth trajectory is related to maternal-fetal nutrition exchange and maternal health[119]. It is important to understand normal fetal growth variability in order to differentiate between baby's who are constitutionally small but grown to their highest potential and those who are small due to adverse intrauterine conditions[119].

Small for gestational age (SGA) is the neonatal outcome that combines the information of birthweight and gestational duration. SGA are children whose birthweight is considered too small than expected with reference to a given growth curve for pregnancies without complications. Still, there is no consensus on the growth curve, and three classification-methods are

commonly used in Sweden; ultrasound-based[120], customized[121], and population-based[116].

Ultrasound-based growth curves were based on the study of fetal weights estimated based on ultrasound measurements of biparietal diameter, abdominal diameter, and femur length. Measures were taken between 76th and 286th gestational day. Four-degree polynomials were used to fit the data of 759 fetal weights for gestational age. The growth curves were established separately for female and male fetuses. In customized growth curves, except fetal sex, maternal characteristics are included, such as maternal height or weight, parity, and ethnicity[121]. Population-based growth curves are based on the Skjaerven's expected birthweights estimated for the Norwegian population[116].

The strictest SGA-classification is based on ultrasound-based definition. Only children with birthweight above two standard deviations with ultrasound-based growth curve as reference are classified as SGA. For population-based or customized growth curves, children with birthweights below the 10th percentile are classified as SGA. Between 2010 and 2012 in Sweden, there were 2.2%, 10%, and 17% of children classified as SGA according to ultrasound-based, population-based, and customized definitions (Figure 7).

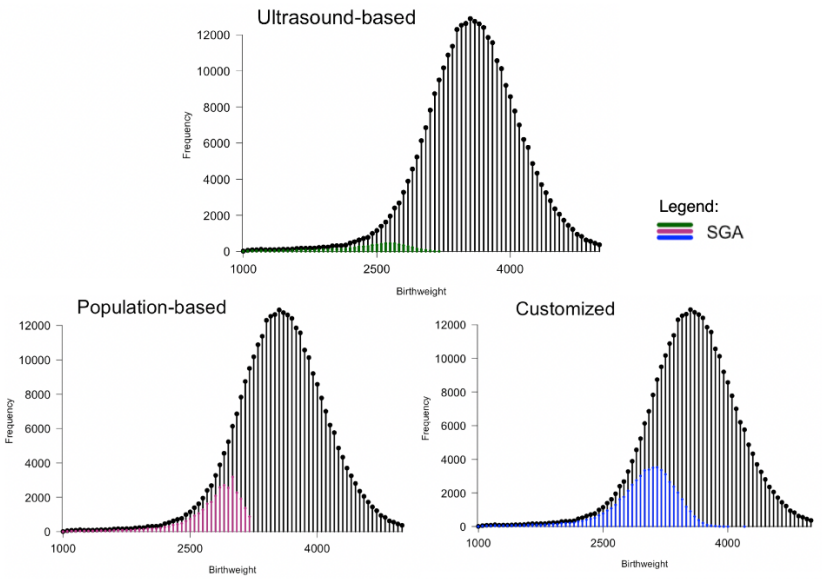


Figure 7. Birthweight distributions with marked proportions of birthweights classified as small for gestational age, Swedish Medical Birth Registry 2010-2012, number of births: 318 373.

04 DATA AND STUDY POPULATIOS

4 Data and study populations

4.1 Introduction

The projects are based on two Scandinavian populations, Swedish (Project 1) and Norwegian (Project 2). The advantage of doing research on Scandinavian populations is the availability of already collected phenotype-rich data. Both Sweden and Norway have multiple national registries established as early as 1947[122]. Personal information is collected throughout the life of an individual and gathered in different registries. In each registry, an individual might be identified based on the unique national personal identification number assigned to each resident. That personal number allows to link different registries accurately and gather comprehensive data for a particular person. That creates unique possibilities for epidemiological research.

4.2 Project 1 (Papers I and II)

In this project, we used two national registries, Swedish Medical Birth Registry (MBR) and Swedish Multi-generation Register.

MBR was established in 1973 with the purpose of collecting pregnancy-related information for public health surveillance. MBR is maintained by the Swedish National Board of Health and Welfare. Registry collects information on all pregnancies occurring in Sweden. Every year, between 1 to 4 % of pregnancies are missed[123]. According to the last quality report on the MBR available from 2003, missing data are considered random, not affecting the association estimates[124].

In MBR, the sources of data, as well as contents, were changing over the years. In the first years of the registry (1973-1982), data from delivery units were the only source. Later, information from antenatal-care clinics (i.e., medical care during pregnancy period) and pediatric wards were included. Nowadays, MBR includes information on maternal demographics, maternal and fetal diagnosis, delivery mode, information on the previous pregnancies.

The Swedish Multi-generation Register is another national database established in 1947 when personal identification number was implemented. It is maintained by Statistics Sweden, and it covers all Swedish population. Initially, the registry was collecting information on the biological parents, and later it was extended to information on adoptive parents, date of adoption, year and country of birth of biological/adoptive parents, date of immigration to Sweden. In the registry, paternity was based on a marital data of a woman (a husband is registered as a father of a child), and in other cases, paternity is based on the acknowledgment or established by a court.

4.3 Project 2 (Papers III and IV)

In this project, we used the study cohort and national registry, the Norwegian Mother, Father, and Child Cohort Study (MoBa), and the Medical Birth Registry of Norway (MBRN).

MoBa is a study aimed to bring insights into the etiology of pregnancy-related outcomes. The study is conducted by the Norwegian Institute of Public Health. For its purpose, pregnant women from all over Norway were recruited. Participation was voluntary, and women were enrolled between 1999 and 2008. At a participation rate of 41%, the final cohort included 95 200 pregnant women. The information on potential determinants or risk factors for pregnancy outcomes was obtained by distributing the questionnaires to the participants. The questionnaires were handed over at multiple gestational ages. Each questionnaire was targeting a specific topic, including a woman's life condition (marital status, education, earnings, etc.) diet habits. Such an approach allowed to collect comprehensive information covering a broad range of exposures.

As studied by Nilsen et al.[125], women included in the MoBa cohort differed from the average woman who gave birth in Norway; they are a sub-population not representing the Norwegian population profile accurately[125]. The ones who agreed to participate in the study were more health-oriented than the general population. In the MoBa cohort, there was a higher percentage of non-smoking women, using supplements, older, and having their first child, in comparison to the pregnant Norwegian population[125]. Thus, as it might be expected, adverse outcomes such as stillbirths or neonatal deaths are less prevalent in the MoBa cohort[35]. Such sub-selection might raise questions about

whether the epidemiologic analyses performed on the cohort are subjected to bias. Women's participation might be related to some characteristics that are associated with both exposures and outcomes. That relationship might result in the association between some exposure and outcome, but in such a case, the association is statistical, not biological. Nilsen et al. compared the associations between 8 pairs of exposures and outcomes estimated in the MoBa cohort and MBRN[125]. They found no statistical difference between the association estimates, but still, they did not exclude the possibility of bias occurring.

The national registry used in this project, MBRN, was established after the world-scale disaster caused by the drug thalidomide[126,127]. Produced by a German pharmaceutical company, the thalidomide was prescribed to a pregnant woman resulting in birth defects such as limb deformities. This catastrophe changed the drug development process[128] and motivated higher epidemiological surveillance at national levels worldwide[126]. Therefore, the initial objective of MBRN was to detect and prevent pregnancy-related outbreaks and threats in the Norwegian population[129].

MBRN, initiated by the Central Bureau of Statistics in 1967, today is run by the University of Bergen, maintained by the National Institute of Public Health, and owned by the State Health Inspectorate[126]. Analogically to Swedish MBR, MBRN collects the data from antenatal units, maternity, and neonatal departments[126]. Content of the MBR was adjusted and changed over the years.

05 PROJECT 1

5 Project 1 – Importance of genetics for gestational duration

5.1 Background

In the field of heritability of gestational duration, there are seven leading studies that vary in the reported estimate (Figure 8). The overall heritability estimate varies between 6 and 36 percent (Figure 8)[130–132]. Differences in the estimates can be largely attributed to different study populations used in the studies, showing already the impact that these can have on the estimate of interest.

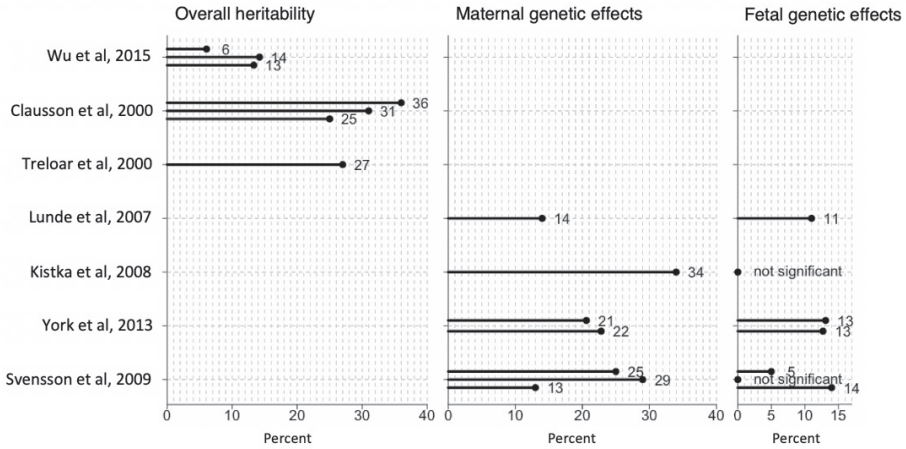


Figure 8. Variability in the estimates of heritability, maternal and fetal genetic effects in the literature.

Wu et al.[130] obtained the lowest heritability estimates, between 6 and 14 percent. The variability in the heritability estimates is due to retrieving the estimate from different family pairs: father-child, heritability [h^2] = 6.77%; mother-child, h^2 = 14.21%; siblings, h^2 = 13.3%. In the analyses, deliveries were restricted to spontaneous (excluded, congenital anomalies and high-risk obstetric conditions in mothers, sample size ~ 1 800 00 births). Gestational duration was based on the LMP date.

In Wu et al.[130] study, father-child pairs served to estimate heritability, while in the study of Lunde et al. the same pairs of relatives served to estimate

fetal genetic effects (Table 2). Both Wu et al.[130] and Lunde et al.[133] observed higher correlation estimates between mother-child than father-child (Table 2). Both assumed that the correlations are the results of common genetic factors and not environmental[130,133]. However, they differed in the interpretation of those correlations. Wu et al.[130] expected to observe high correlation estimates between both parents and their offspring according to the high genetic relatedness between them (50%)[10]. The observed correlation was lower-than-expected due to the contribution of environmental factors. Wu et al.[130] explained that heritability estimates based on parent-offspring pairs are biased by intergenerational factors. Lunde et al.[133] had a different perspective: they did not focus on relatively small estimates of parent-offspring pairs, but they considered the difference between correlation estimates of mother-child and father-child pairs. Lunde et al.[133] explained that higher estimates for mother-child are due to the underlying additive effect of fetal and maternal genetics while father-child pairs represent the effect of transmitted parental genes to fetus only (fetal genetics).

Studies family pairs	**Genetic effects - interpretation**		**Correlation/regression estimates**	
	Wu et al.	Lunde et al.	Wu et al.	Lunde et al.
Father-child	½ additive genetics	½ fetal genetics	0.03	0.06
Mother-child	½ additive genetics	½ fetal genetics + ½ maternal genetics	0.07	0.13
Full siblings	½ additive genetics + ¼ dominance + common environment	½ fetal genetics + maternal genetics + common environment	0.25	0.32
Maternal half-siblings	¼ additive genetics + common environment	¼ fetal genetics + maternal genetics + common environment	0.03	0.22
Paternal half-siblings	¼ additive genetics	Not studied	0.19	Not studied

Table 2. Summary of the studies of Wu et al.[130] and Lunde et al.[133].

Analogically to Wu et al.[130], Clausson et al.[131] dismissed the use of intergenerational cohorts to estimate heritability as it may be affected by changes in environmental conditions. Instead, they used female twins and their

offspring. The offspring of female twins were born between 1973 and 1993 (Figure 9). Both UL and LMP were used for gestational duration estimation. Clausson et al.[131] provided estimates based on subsets of primipara and multipara twin mothers. In the case of multipara mothers, the mean of gestational ages at birth of all children was taken into account in heritability analysis. Furthermore, Clausson et al.[131] estimated heritability using dichotomized (gestational duration < 37, PTD) and continuous gestational duration. They obtained heritability estimates of 36% for PTD, 31% for gestational duration calculated on mothers restricted to primipara, and 31% when all children (the mean) of a mother were considered (Table 3).

Studied family pairs	**Clausson et al.**		**PTD (<37)**	**GA**
	Genetic effects	Children considered	**Correlation estimate**	
Children of monozygotic twin mothers	additive genetics + shared environment	1st child	Not studied	0.25
		The mean of all children	0.40	0.31
Children of dizygotic twin mothers	½ additive genetics + shared environment	1st child	Not studied	0.11
		The mean of all children	0.09	0.13
			Heritability	
		1st child	Not studied	0.25
		The mean of all children	0.36	0.31

Table 3. Summary of Clausson et al.[131] study.

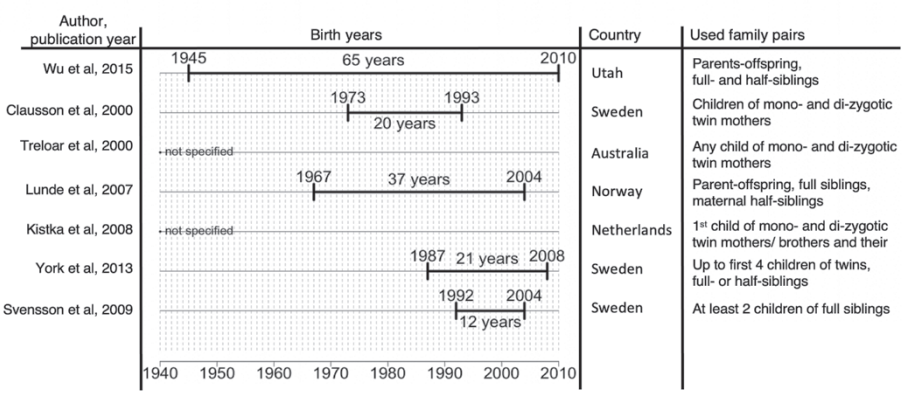


Figure 9. Birthyears ranges of the relatives included in the analysis of heritability in the literature.

Following Lunde et al.[133] interpretation of the correlation, twin sisters might be similar in their gestational durations only through their shared genetics, whether expressed through the mother's ability to hold pregnancy or through genes transmitted to the fetus. In such perspective, Clausson et al.[131] reported heritability estimates represent maternal genetic effects. Kistka et al.[134] found no heritability when performing the analysis based on male-twins (monozygotic and dizygotic) and 34% when performed on the female-twins (analyses restricted to their first child). Analogically to Lunde et al.[133], the difference in the estimates based on male- and female-twins Kistka et al.[134] interpreted as only maternal genetic effects and no fetal contribution to gestational duration. Two other studies indicated higher maternal genetic effects than fetal, regardless of the sample with the exemption of the analysis performed on iatrogenic deliveries. Svensson et al.[135] reported a heritability estimate of 14% for fetal and 13% for maternal genetic effects when restricting the sample to iatrogenic deliveries.

The estimates of genetic contributions are retrieved under the assumptions of no genetic and environmental interactions, constant genetic and environmental effects, random mating, and no transmission of life-style factors[10,133,136]. The assumption of constant environmental factors was observed to not hold among intergenerational cohorts such as parents-offspring. Parents and their offspring are born in distant periods of time characterized by different environmental conditions affecting the variability of gestational duration. As an example, in the cohorts of parents, there are fewer cases of PTD than in the cohorts of their offspring. In part, this might be due to lower survival chances of prematurely born children in earlier years. Also, the number of inductions and cesarean sections increased over the years, increasing the differences in factors affecting the variability of gestational duration between parental and children cohorts. Thus, the parental cohorts were dismissed as introducing bias attributed to time-related environmental factors[130,131,137].

Time-related environmental factors might not only introduce differences among intergenerational pairs but also among children of siblings, including twins. The children of paired relatives might be born in the periods of time characterized by different obstetrical practices (Figure 9). Therefore, time-related environmental factors might also contribute to the association estimates between pairs other than parents-offspring. It should be noticed that in the

characterized by different obstetrical practices (Figure 9). Therefore, time-related environmental factors might also contribute to the association estimates between pairs other than parents-offspring. It should be noticed that in the presented studies on the heritability of gestational duration, the analyses were not controlled for birth year[130–136].

5.2 Aim

In Paper I, we analyzed whether time-related environmental factors are present and contribute to the association between full- and half-siblings. In Paper II, we assessed the contribution of factors related to gestational duration estimation or reporting on gestational duration variability.

5.3 Methodological aspects

QQ-plots were used to detect differences between distributions of gestational duration in the cohorts of paired relatives; parent-offspring, full-, and half-siblings. Next, correlations between those pairs were estimated. For the association analyses, gestational duration was adjusted for maternal age and parity. The analyses were performed in different subgroups according to the age-gap categories between relatives. The stratification was used to detect the presence of time-related environmental factors, even for non-intergenerational relatives.

For defining time-related environmental factors in the Swedish population, we studied the distribution of gestational duration. We analyzed temporal changes in distribution's shape and in the interval between LMP-date and due date.

5.4 Results and discussion

The distributions of full siblings differed regardless of the age gap between them. A higher rate of PTD was observed in the cohort of earlier born siblings if the age gap was lower than seven years (Figure 10, left panel). This pattern was reversed if we restricted the graphical analysis to spontaneously delivered neonatal period survivors (Figure 10, right panel); higher PTD rates were observed in the later-born cohorts. Following the reasoning presented in the literature[130], the QQ-plots of full siblings showed that there are environmental

factors contributing to the differences in the duration of pregnancies of a single woman. It is possible that PTD in one pregnancy might motivate for more antenatal visits or adjust her lifestyle in the next pregnancy. Also, a woman's body undergoes changes during the pregnancy that might affect the outcomes in the next pregnancy. After uterine vascular remodeling, the trophoblast invasion thus, placentation in the next pregnancy is easier. Better placentation results in better nutrient flow. Besides the flow of nutrients, there is also cell exchange between the mother and fetus, which is progressing with gestational age (phenomena called "microchimerism"). Fetal cells might be found in various maternal tissues, including the liver, lymph nodes, intestines, lungs, skin, or spinal cord[138]. They remain in a woman's body for decades after giving birth. In the next pregnancy, the mother transfers her own and previous child cells to a new fetus[138]. Such physiological changes create that assumption in heritability analysis of constant environment between pregnancies of a woman does not hold.

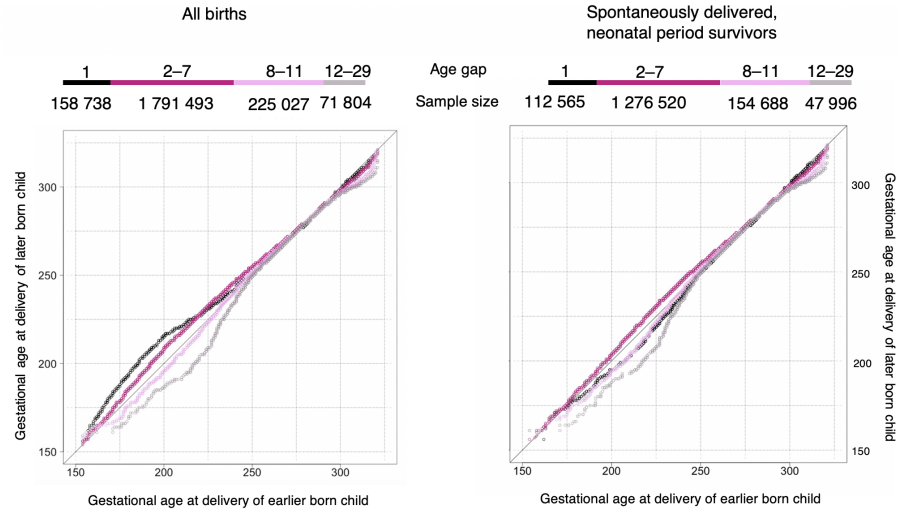
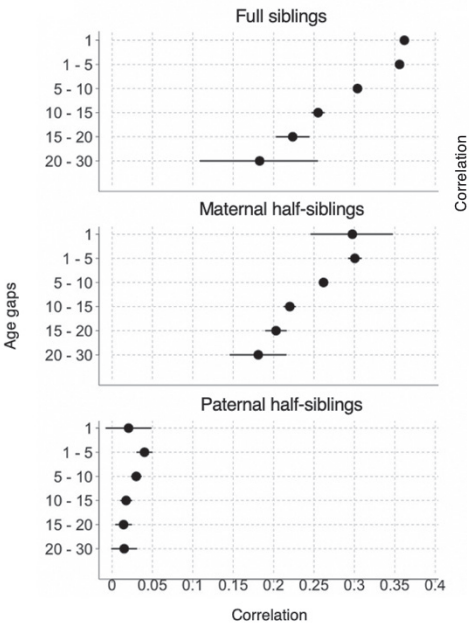


Figure 10. QQ-plots between the cohorts of older and younger full siblings. Left panels: all births from the Swedish Medical Birth Registry 1973-2012, right panel: cohort restricted to spontaneously delivered neonatal period survivors.

The lack of constant environmental conditions could also be supported by the correlation analyses. Correlation estimates were decreasing with an increasing age gap between full siblings and maternal half-siblings (Figure 11). Furthermore, the correlation between full siblings born at different hospitals have lower correlation estimates than siblings born in the same. Different

hospitals might report or estimate gestational duration differently, contributing to lower correlation estimates. Also, the correlation between full siblings born a maximum of two years apart increased over the years (Figure 12). Differences in obstetrical practices over the years might underlie the correlation trend among full siblings with a small age gap. Standardization of the estimation methods and obstetrics practices might lower the variability of gestational duration and increase the association estimates between relatives.



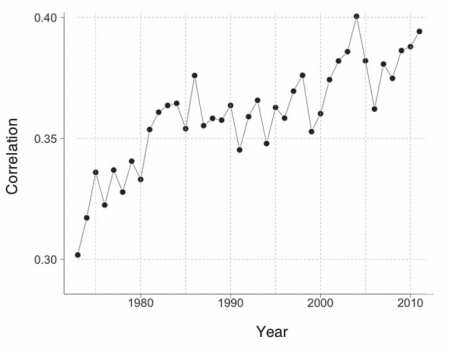


Figure 12. Correlation between full siblings born maximum two years apart over the years.

Figure 11. Correlation between siblings with regard to their age gaps.

Despite that distributions between paired relatives differed in a similar manner, the correlation estimates varied. For example, mother-child and maternal half-siblings had similar QQ-plots but different correlation estimates (Figure 13). In heritability analysis, correlation between mother-offspring pairs is assumed to be due to the sum of half fetal- and half maternal genetic effects, while the correlation between maternal half-siblings is due to the sum of a quarter of fetal- and whole maternal genetic effects and shared environmental conditions[10,133]. Maternal half-siblings may share more environment compared to mother-child pairs, creating higher correlation estimates.

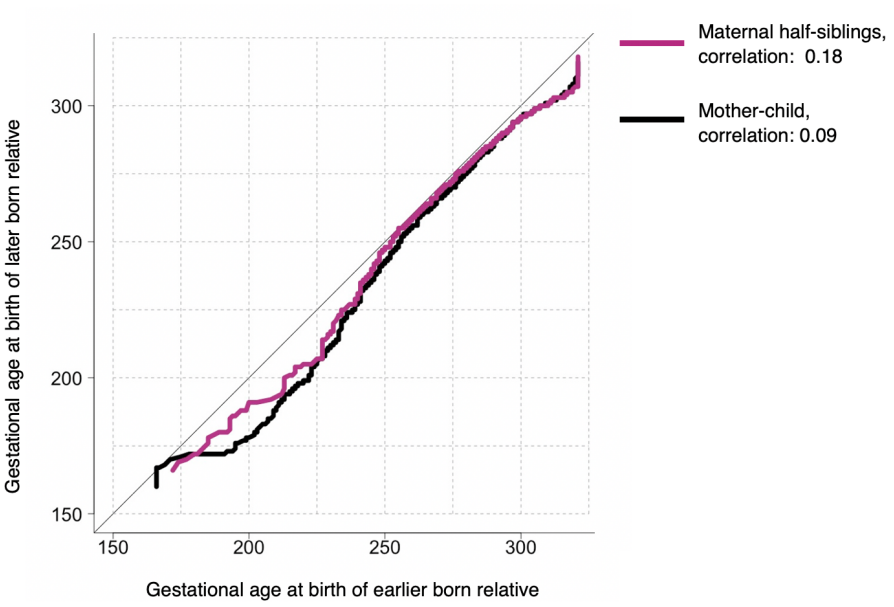


Figure 13. QQ-plots and correlation estimates of relatives: mother-child and maternal half-siblings.

The distribution of gestational duration was constantly changing over the years. Those changes are related to different approaches taken to estimate or report gestational duration. Among the most acknowledged is the shift of the whole distribution towards lower values of gestational duration. The shift was due to the change in the estimation method: from LMP- to ultrasound-based. Subtle and less noted sources of gestational duration variability also exist and belong to changes in reporting. Before around 1995, it was frequent to report gestational duration in weeks instead of days. The estimates transferred to the Swedish Registry were converted into days by multiplication by seven and adding three (half week). This reduced the natural variability in gestational durations and artificially created peaks at some durations: these were more frequently reported (every seventh day starting from 157 day) (Figure 14).

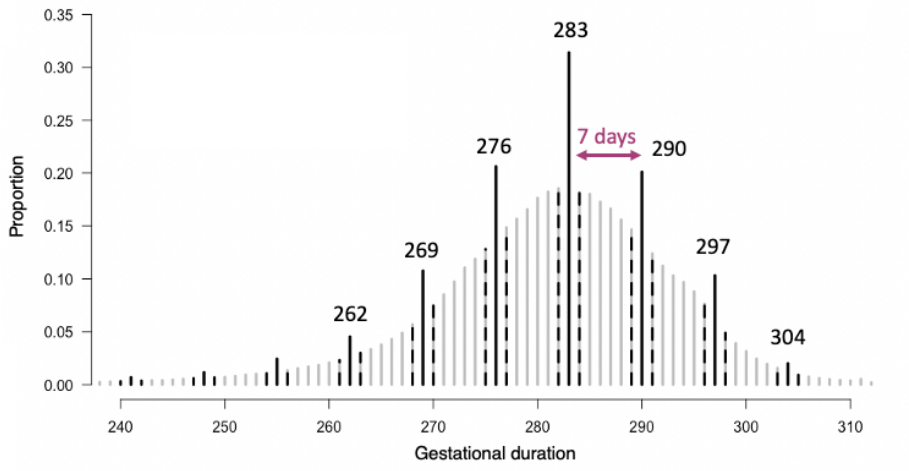


Figure 14. Gestational duration distribution, the Swedish Medical Birth Registry, 1973-1990.

Analyzing the median value of the interval between LMP-date and LMP due date depicted the changes in the assumed expected gestational duration and gestational duration variability related to leap years. Plotting the median over the years showed that before 1990, the common expected gestational duration was 280 days and later 279 (Figure 15). The expected gestational duration is used to define the due date, and the due date is, in turn, used to estimate gestational duration. Hence, variability in the expected gestational duration contributes to the variability in gestational duration estimate. Furthermore, until 2004, the interval between LMP-date and LMP due date was commonly one day longer every leap year in comparison to neighboring years (Figure 15). A bigger interval was the result of not accounting for an extra 29th day in February of leap years when estimating the due date. Again, it added noise to the gestational duration variability.

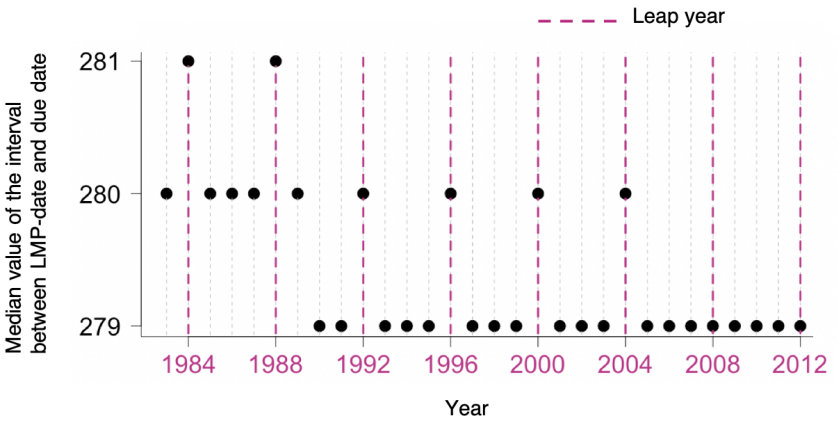


Figure 15. The median value of the interval between last menstrual period day and due date over the years.

Plotting the median interval between LMP-date and LMP due date over the month and year in which a woman gave birth allows to see the above-mentioned larger variability before 1990 (Figure 16). Women who gave birth between December and February tend to have longer intervals than women delivering in other months. The variability resulted from estimating a due date by adding expected gestational duration in months (+days) units and not days.

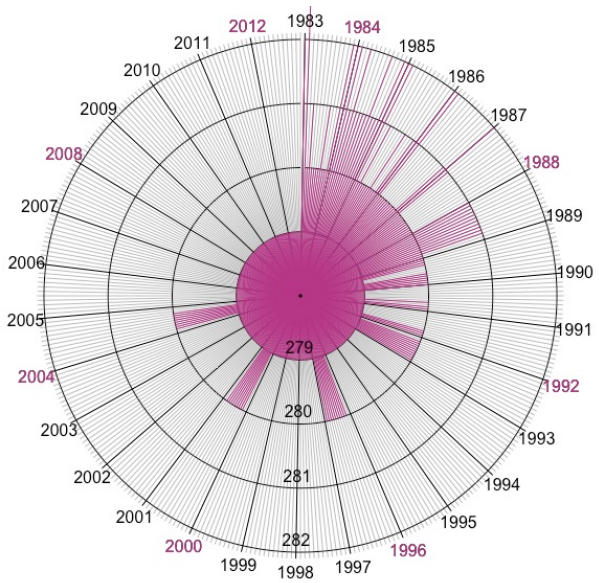


Figure 16. The median value of the interval between the last menstrual period day and the due date for women giving birth in given month and year.

5.5 Summary

In this project, we observed variability in the association in gestational age at delivery between relatives with respect to their birth years or age gaps. Such observation suggests the presence of temporally changing environmental factors even between relatives born or delivering within a small-time frame. Next, we showed that data handling or obstetrical practices change over time, contributing to the gestational duration variability. Heritability estimates, which are designed to assess the importance of genetics for any trait, rely on the assumption of constant environmental conditions. Two studies included in this thesis suggest that the assumption might not hold.

Variability in the heritability estimate might be explained by the complex combination of factors constructing correlation estimates. Simplified covariance composition underlying heritability analysis might not be valid in normal life conditions, leading to biased heritability estimates and missing heritability. Even if genetically identical, relatives might differ if affected by the environment. For example, monozygotic twins (i.e., twins having the same genome) are born already with different fingerprints and iris patterns all due to differences in their uterus' microenvironments[139].

In research, heritability remains to be one of the controversial concepts. The concept involves the idea that organism's development might be predicted by understanding separately the average effects of genes or environments. It opens a debate on what is more important for a trait development, nature (genes) or nurture (environment). That has been a source of controversy for a long time.

In the 1970s, one of the biggest heritability opponents, Richard Lewontin, emphasized that the nature-nurture split arose from the misunderstanding of the causes of variability in a phenotype[140]. The alterations in a phenotype might be due to a single gene effect when a phenotype is a result of homozygosity for a rare deleterious gene or due to many genes with small effects. Lewontin[140] suggested that misunderstanding lies in associating the importance of the environment only for the second cause, while the interactions between genotypes and environment are present in both cases. Such one-sided association led to erroneous perception that phenotypic variability is due to

environmental or genetic forces, while variability in a phenotype arises from the two alternative genetic bases.

Lewontin[140] suggested the misconception introduced with the heritability concept. The heritability concept was used to discuss the causes for phenotypic variability, but in fact, it relates to the components underlying every phenotype. Furthermore, he pointed out that the misconception was maintained and strengthened by misunderstanding the analysis of variance. The heritability concept was supported by presenting analysis of variance as a tool to draw conclusions about the causality. Lewontin[140] argued that analysis of variance is unable to differentiate the reasons for the estimate of genetic variance. Small genetic variance might be due to a lack of functional relationship between the genotype and phenotype or due to population homogeneity for the specific gene of interest.

Arguments against the heritability concept were dismissed as driven by the fear of potential social and political consequences of such dichotomization and not scientific exploration[12]. Instead, the concepts used in heritability analysis were used to claim the legitimacy and relevance of the analysis.

06 PROJECT 2

6 Project 2 – Maternal caffeine/selenium intake and neonatal outcomes

6.1 Background

When studying the effect of prenatal exposures on neonatal outcomes, often birthweight and gestational duration are taken into account[141]. Birthweight and gestational duration are included in regressions with the aim of removing potential bias attributed to confounding or to improve the precision of the estimates[141]. However, as suggested by Hernandez-Diaz et al.[5] such model adjustment is inappropriate as both birthweight and gestational duration are not confounders but potential predictors.

By definition, a confounder is a variable that affects both exposure and outcome, and a predictor is the factor affecting the outcome. Birthweight and gestational duration are occurring after the exposure; thus, they cannot affect it, but both are considered as strong determinants for neonatal outcomes[115]. Hernandez-Diaz et al.[5] suggested that adjusting for predictors is inappropriate when the aim of a study is to assess the overall effect of prenatal exposure on neonatal outcomes.

Hernandez-Diaz et al.[5] pointed that by adjusting the exposure-outcome model for birthweight or gestational duration, we enter into the assessment of causal rather than overall association. If the exposure is associated with birthweight and birthweight with a specific neonatal outcome, the observed association between the exposure and the neonatal outcome might be due to its effect on birthweight. In such a case, the effect of the exposure on the outcome is called indirect (Figure 17). An indirect effect might be observed if the adjustment for birthweight in the model of exposure on outcome results in a lack of association (Figure 17, column 'Expected odds ratio'). If the association after birthweight adjustment remains significant but differs from the unadjusted model, it is interpreted as the exposure has a direct effect on the outcome.

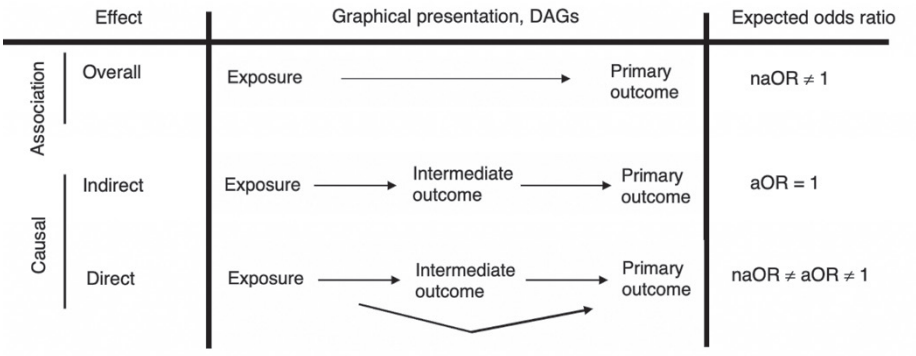


Figure 17. The concept of overall and causal associations. Abbreviations: directed acyclic graphs, DAG; unadjusted odds ratio, naOR; adjusted odds ratio, aOR

In our research group, there was an interest in the effect of maternal nutrition on perinatal outcomes. Sengpiel et al.[1] found an association between caffeine intake and SGA, and Barman et al.[2] reported a relationship between selenium intake and PTD, and Sole-Navais[3] between selenium intake and SGA. Both SGA and PTD are associated with neonatal outcomes. Therefore, maternal caffeine or selenium intake were expected to be at least indirectly associated with neonatal outcomes.

6.2 Aim

The aim of the two papers was to assess the association between maternal consumption of caffeine (Paper III) and selenium (Paper IV) on neonatal outcomes. The results were discussed in the perspective of Hernandez-Diaz et al.[5] methodological considerations.

6.3 Methodological aspects

6.3.1 Exposures

Maternal intakes of caffeine and selenium were estimated based on maternal self-reported diet habits (covering pre- and pregnancy periods). Diet habits were reported once in a questionnaire handed over to women in week 22 of gestation.

The effect of caffeine was assessed based on the continuous variable of total or source-specific caffeine intake (sources: coffee, energy drinks, tea,

chocolate) (Figure 18). Selenium effect was assessed based on the continuous selenium intake from diet or supplements (inorganic or organic) (Figure 18). Total selenium intake (sum of the intake from diet and supplements) was dichotomized and presented in the form of the two exposures: selenium intake < 30 µg/day and selenium intake > 120 µg/day (Figure 18).

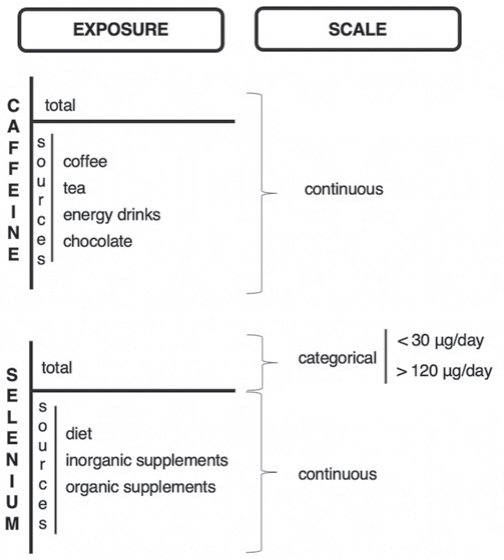


Figure 18. Graphical presentation of the exposure variables.

6.3.2 Neonatal outcomes

The health condition of a newborn might be represented in two ways, by the ICD-code reflecting specific health issues or by the medical intervention the baby has received after birth. In such a way, we created two complex variables that we called 'neonatal mortality/morbidity' and 'neonatal intervention.'. To represent poor neonatal health, to the neonatal mortality/morbidity we included the most frequently occurring adverse health events (Table 4), and to medical intervention, admission to neonatal intensive care units, or application of respiratory or antibiotics treatments.

Category		Specific morbidities (ICD-10)
Vascular issues	Reduced oxygen and nutrient transport to the brain and other organs	Birth asphyxia (P21), other disturbances in the cerebral status of the newborn (P910–916)
	Bleeding in the brain caused by rupture or leak of vessels	Intracranial (non-traumatic) hemorrhage of the fetus and newborn

	Abnormal vessels development in the eye leading to blindness	Retinopathy of prematurity (H351)
Intestinal diseases		Meconium ileus/necrotizing enterocolitis (P75, P76, P77, P780, P781)
Infections		Bacterial sepsis in the newborn (P36), other infections occurring in the perinatal period (P39)
Respiratory diseases		Chronic respiratory disease originating in the perinatal period (P27)
Neonatal death		

Table 4. Morbidities included to the variable neonatal mortality/morbidity.

6.3.3 Cohort

In both studies on caffeine and selenium, we included only singleton live-born babies. In the caffeine study, the cohort was additionally restricted to term-born babies. Data cleansing involved removal of women with invalid energy intakes (Paper III and IV), missing dietary reports (Paper III and IV), questionable gestational duration (< 22+0 and > 41+6 weeks), and outliers for birthweight or selenium intake (> 4 standard deviations, Paper IV).

6.3.4 Data analyses

Logistic regression was used to assess the association between food components and neonatal outcomes. Both studies used a similar set of variables as confounders (Table 5).

Confounders		**Caffeine study**	**Selenium study (Paper IV)**	
Categories	**Specific characteristics**	**(Paper III)**	Selenium from diet or supplements	Whole-blood concentration
	Maternal BMI	X	X	X

Maternal basic characteristic	Maternal education	X	X	X
	Household income	X	-	-
	Marital status	X	-	-
	Maternal age at delivery	X	X	X
Substance abuse	Smoking during pregnancy	X	X	X
	Passive smoking	-	X	X
	Alcohol consumption during pregnancy	X	-	-
Pregnancy-related	Nausea in 22nd week of gestation	X	X	-
	Parity	X	X	X
	Planned pregnancy	X	-	-
	Baby's sex	X	-	-
Diet-related	Dietary fiber intake	-	X	-
	Iodine intake	-	X	-
	Protein intake	-	X	-
	Total energy	X	X	-
	n-3 intake	-	X	-

	Folic acid supplementation	X	-	-
	Other sources of the exposure	-	X	-

Table 5. Confounders included in the models studying the effect of caffeine, selenium from diet and supplements, or selenium whole-blood concentration.

6.4 Results and discussion

After controlling for the set of confounders, there was no association between total caffeine or selenium intake and neonatal outcomes when selenium was considered at the continuous scale (Figure 19). Only maternal caffeine intake from chocolate was associated with neonatal intervention (Figure 19). This association might be due to unobserved confounding factors. Increasing chocolate intake might be associated with other unreported diet habits that are associated with neonatal outcomes (snacking, sweets consumption). Low maternal selenium intake (< 30 µg/day) was associated with both neonatal outcomes. Since selenium is an essential nutrient protecting cells against free radicals, the association could indicate the importance of this element for pregnancy and a child's health.

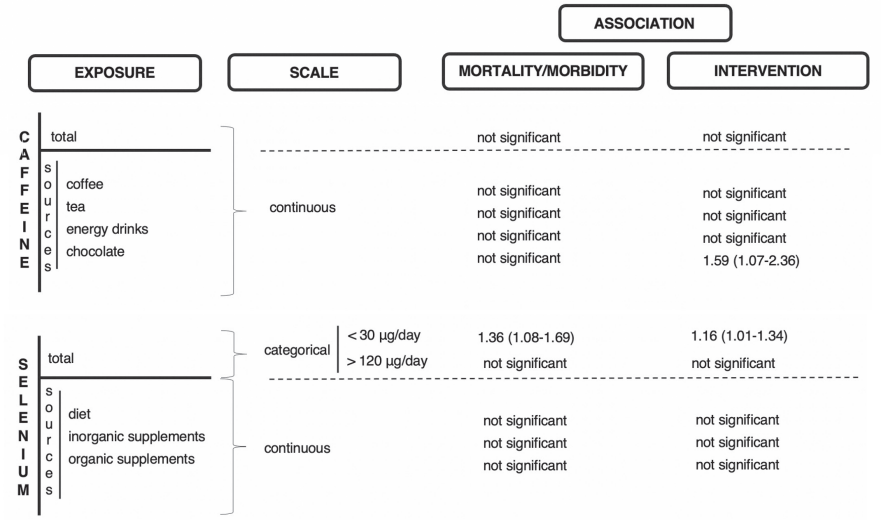


Figure 19. Odds ratio of neonatal outcomes (mortality/morbidity or intervention) for different exposures in the adjusted models. 'Not significant' indicates statistically insignificant odds ratios.

Given the intermediate associations (caffeine or selenium intake with SGA or PTD in the case of selenium) and SGA with neonatal outcomes, the lack of association in this study was unexpected. We anticipated that the food components we studied would be associated with neonatal outcomes indirectly through SGA (or PTD), regardless of whether the intermediate associations were causal or due to residual confounding (Figure 20). Additional adjustment for SGA (in the caffeine model) or SGA and PTD (in the selenium model) did not affect the results (Figure 21). A possible explanation is that SGA babies vary in their health condition due to the underlying cause leading to SGA (Figure 20). Also, SGA might not be a determinant of neonatal outcome but a feature indicating a condition of disease (a symptom) (Figure 20).

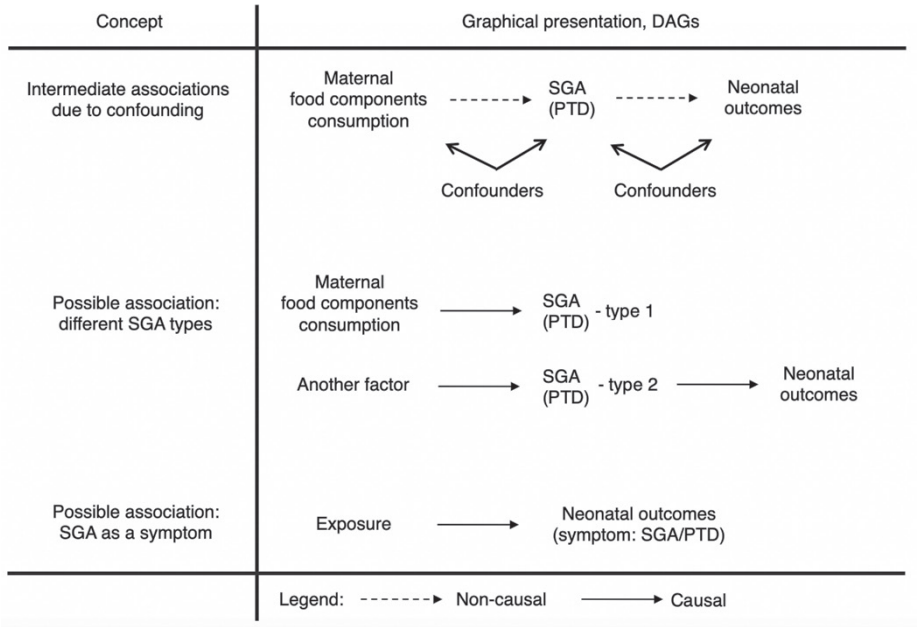


Figure 20. Possible associations between maternal food components, SGA, and neonatal outcomes. Abbreviations: SGA, small for gestational age; PTD, preterm delivery; directed acyclic graphs, DAG.

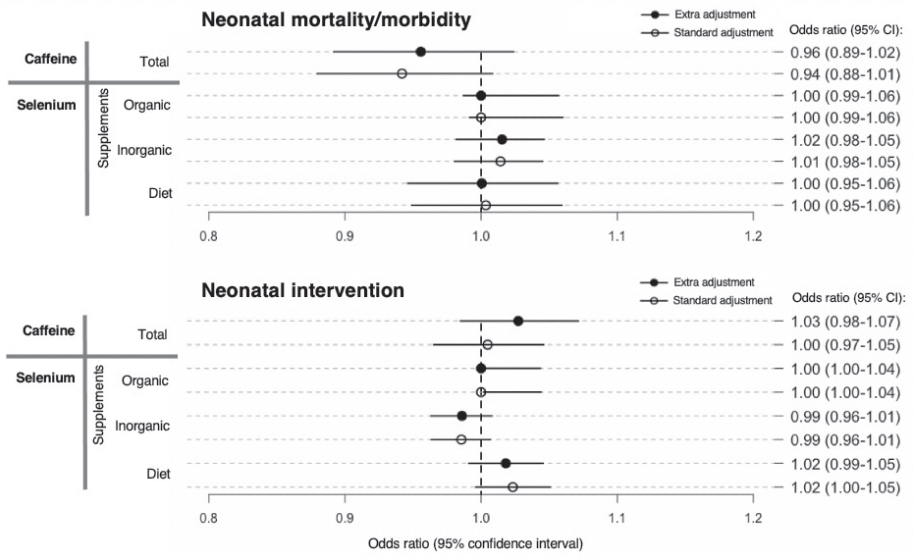


Figure 21. Odds ratio of neonatal outcomes (mortality/morbidity or intervention) for continuous caffeine and selenium intakes. Two types of models were presented: adjusted for standard set of confounders (open dot) and additionally adjusted for SGA (in caffeine models) or SGA and PTD (in selenium models).

6.5 Summary

Methodological consideration of Hernandez-Diaz et al.[5] relates to the problem of model specification. Model specification should be considered according to the goal of the study. We might be interested in the overall association for predictive purposes that might be of interest to both clinical and public health communities. Understanding whether the association is due to causal effects (unbiased) might be useful to build effective interventions. In exploring causal associations, we might aim to just detect causal factors (assessing the significance), or we might be interested in understanding the change they introduce in the studied outcome (assessing the size effect).

Incorrect model specification might introduce bias to the estimates or reduce its precision. An unbiased estimate is the one that would be equal to the real value if there are no random effects. Estimate's precision is represented by its standard error or confidence interval; the higher the precision, the lower the standard error. Bias and precision of the estimate are differently affected in different models. Hernandez-Diaz et al.[5] pointed at the importance of both confounders and predictors in the logistic models. The estimates of logistic

regression are sensitive to the presence/absence of confounders[142]. In fact, these are also sensitive to predictors. The nature of logistic regression makes accurate estimates more difficult to obtain when the outcome is affected by a broad set of different factors. The multifactorial background is anticipated in the case of a majority of perinatal or neonatal outcomes. Furthermore, as the outcome of interest in different populations or groups of individuals might be affected by different sets of factors, comparing the odds ratios from different strata might lead to incorrect conclusions about the association between a given exposure and the outcome[143]. In contrast to the logistic regression, coefficients in the linear regression are affected only by confounders[142]. Such property supports the use of perinatal outcomes on a continuous scale, where possible.

07 SUMMARY

7 Summary

Epidemiological research on pregnancy outcomes is challenging due to the complexity of pregnancy. Pregnancy, like no other medical subject, includes two individuals, the fetus and mother. Perinatal outcomes are the results of the interplay of signals between them. The normal chain of reactions might be affected by multiple factors over the pregnancy. The sensitivity of the pregnancy varies over time and relates to the development stage of the fetus and placenta. Furthermore, every next pregnancy might differ from the previous; primiparity alone is considered as a risk factor for PTD. The complexity of pregnancy gives rise to multiple questions on how to perform epidemiological research to obtain unbiased estimates. How should we assess the familial aggregation of a trait? Should we take the mean of all pregnancies of a woman, pick one gestation of a specific parity, or consider multilevel modeling? What can we gain or lose when we dichotomize the outcomes? The answer depends on the scientific question. The projects in this thesis relate to the two questions: 1) is it feasible to assess the extent to which genetics contribute to the variability, 2) how does the model adjustment affect our estimates?

In the first project, we aimed to contribute to the reasons for the occurrence of missing heritability. Commonly, the occurrence of this phenomenon is attributed to the limitations of GWA studies. In this project, we took a step back to consider the validity of heritability analysis. Estimates of heritability are established for a population in very specific conditions: Hardy-Weinberg equilibrium, mating at random, allele frequencies equal in the sexes, no migration, no mutation, and constant environment between compared relatives. In this thesis, we considered the last assumption, constant environment.

The assumption of constant environment is important because it underlies assumed composition for the covariance between the relatives. For example, the covariance between parents-offspring pairs is defined as half additive variance, while between full siblings as the combination of additive variance, dominance effect, and shared environments. In the studies included in this thesis, we reported temporal variability in the relationships between gestational

ages at birth of different relatives. This variability suggests the presence of changing environmental conditions. Also, we presented that gestational duration varies to a substantial extent due to factors related to data management. Differences in the obstetrical practices between health centers and overtime might contribute to the differences between the relatives. Therefore, Project 1 might point at the possible limitations of heritability analysis and motivate further research on its validity.

In the second project, we followed the methodological consideration of Hernandez-Diaz et al.[5], who pointed at the difference between the study of the overall exposure-outcome association and causality. She pointed that if we aim to study the overall association, we might omit the adjustment for predictors in the logistic regression. In our study, we observed no difference between the estimates for caffeine or selenium and neonatal outcomes whether we included or not predictors (SGA or PTD) to the regression. Lack of the changes might suggest diversity in SGA or PTD translated in different neonatal outcomes. Interpreting the results accordingly to Hernandez-Diaz et al.[5] discussion, we assumed that previous associations between maternal diet and SGA were causal or due to confounding. However, we cannot rule out the possibility that associations were due to selection bias. Associations between maternal caffeine intake and SGA were substantial for ultrasound-based definition of SGA (OR = 1.24, 95%CI: 1.16, 1.31) and population-based or customized definitions in the selenium study (OR = 0.92, 95%CI: 0.90, 0.95)[1]. Change of SGA definition resulted in the decreased odds ratio to 1.17 (95%CI: 1.13, 1.20) for population-based and 1.07 (95%CI: 1.04 to 1.11) for customized[1]. The association between maternal selenium intake and SGA based on ultrasound-based definition resulted in a non-statistically significant estimate of OR = 0.93 (95%CI:0.84, 1.04)[3]. Further studies on this topic are required.

08 ETHICAL APPROVALS

8 Ethical approvals

Studies included in this thesis are registry-based; thus, ethics consider mainly data privacy of included individuals. Nordic national registries collect the data with the rule of confidentiality. The information that is stored in the registries is never used for the reasons that could violate the privacy of an individual. Data are released to scientists whose research plan is approved by Ethical Board. Approvals are released to the groups whose research topic and methodology are ethically acceptable and scientifically relevant.

The studies in Project 1 obtained approval from the Regional Ethics Committee of the Western Health Care Region in Sweden (Dnr. 576-13 and Dnr. 091-06). Two datasets were used in Project 1, the Swedish Medical Birth Registry and the Swedish Multi-generation Register. Data were pseudonymized, meaning that the personal identification numbers are known only to data managers, the Swedish National Board of Health and Welfare and Statistics Sweden. Statistics Sweden linked two datasets.

The studies in Project 2 were approved by the Regional Committee for Medical and Health Research Ethics South East in Norway (REK/Sør-Øst 2010/2683). Data in the MoBa cohort were prospectively collected, and participation in the study was preceded by signing informed written consent. The consent applies to biological material, questionnaires, and medical records. The Norwegian Data Inspectorate approved data collection. Participants had the rights to withdraw from the study at any point with the consequence of removing their data from the database.

OTHER CO-AUTHORED PUBLICATIOS

Other co-authored publications

Al-Haddad BJS, Jacobsson B, Chabra S, **Modzelewska D**, Olson EM, Bernier R, Enquobahrie DA, Hagberg H, Östling S, Rajagopal L, Adams Waldorf KM, Sengpiel V. *Long-term Risk of Neuropsychiatric Disease After Exposure to Infection In Utero.* JAMA Psychiatry. 2019 Jun 1;76(6):594-602.

Jacobsson B, Pettersson K, **Modzelewska D**, Abrahamsson T, Bergman L, Håkansson S. *Förtidsbörd största perinatala problemet - 5,7 procent av graviditeter i Sverige slutar för tidigt, inte klarlagt varför – kostar miljardbelopp varje år [Preterm delivery: an overview on epidemiology, pathophysiology and consequences for the individual and the society].* Lakartidningen. 2019 Oct 8;116:FR6F. Swedish.

Sole-Navais P, Bacelis J, Helgeland Ø, **Modzelewska D**, Vaudel M, Flatley C, Andreassen O, Njølstad PR, Muglia LJ, Johansson S, Zhang G, Jacobsson B. *Autozygosity mapping and time-to-spontaneous delivery in Norwegian parent-offspring trios.* Hum Mol Genet. 2021 Feb 4;29(23):3845-3858.

ACKNOWLEDGEMENTS

Acknowledgements

'*It takes two flints to make a fire.*'
- Louisa May Alcott

Big thanks to all the people who were core stones in my PhD journey.

To my main supervisor, **Bo Jacobsson**, for giving me this opportunity to explore more than just the fields of obstetrics or genetics. For giving me space and time to experiment and taste various aspects of research. Thank you for creating an environment to grow, pushing hard forward, showing the things I stubbornly was rejecting to see. For allowing to be me, for your brave, honest and caring attitude you have towards science, patients, and surrounding you world. Thank you for all!

To my co-supervisor, **Staffan Nilsson**, for always cheering up everyone every time, for helping when statistical conventions were leaving me out of breath, for invaluable support in dealing with reviewers' requests, for great jokes out of my mistakes or situational particularities that could be noticed by no one else but you. Your creativity, humor, and cleverness are incredible. Thank you!

To my colleague and supervisor, **Pol Sole-Navais**, for always introducing friendly atmosphere in a group, and being eager to listen to my every controversial observation. For a great support in the last months of my PhD, for cheering up, and always, ALWAYS, being ready to serve and help. Thank you!

To **Tomasz Burzykowski**, some things are hard to be expressed verbally and your person belongs to that list. An extraordinary mixture of high standards, gigantic kindness, concreteness, and such fast facts connection. Thank you for your availability, sharing your great knowledge, experience, and kind support. Thank you so much!

To **Ylva Folkesson**, **Anja Andersson**, and **Kristina Karlsson** for unlimited administrative help, availability, and always sharing kind word. Ylva, for reminding to enjoy every second of PhD journey, and sharing your uplifting perspectives on the things. Thank you!

To **Malin Barman** for your great support in the last struggles in my PhD journey. Thank you for your availability, supervision, and help to push things through!

To **Rino Bellocco** and **Verena Sengpiel** without your support at my very early times of my research work I would have never been here. Thank you for introducing me to the group and the research world. Thank you!

To co-authors, **Anne Lise Brantsæter**, **Helle Margrete Meltzer**, **Christopher Flatley**, **Anders Elfvin**, **Ge Zhang**, **Louis J. Muglia** for your great support and help in the research endeavors. Thank you!

To **Maria Hallingström** for so much warmth and kind support. For sharing your experience, advises, being ready to help, cheering and comforting in the intense moments. Thank you for being a real friend!

To my Best and Dearest Friends, **Mom** and **Dad**, for being my favorite audience, my best advisors and supporters, my biggest blessing and the gift I was bestowed on just at the day of my birth. Thank you for always walking by my side and helping to understand and explore Life. Thank you for always being close when I'm balancing on the edge and reminding me that I can fly when I'm falling.

To my Dzióbeczki Family; **Beata**, **Krzysztof**, and **Emilia Peichert**. For your never-ending support, prayers, gigantic love, laughs, and so much warm feelings. Thank you for always being close even when thousands of kilometers are keeping us apart!

To my Beautiful Grandparents, **Jadwiga** and **Stanisław Drzewieccy**, whose joy on my little achievements I will no longer see. In memory of them, who like no one else were cheerleading my education.

To my great friends, **Magdalena Chomać**, **Sabina Filipovic, Aldona Lindberg**, **Anna Kałużny,** and the whole **Adam Wojnarowski's Warsaw Bible study group**, for your kind hearts, for your words, for your prayers. For overanalyzing every aspect of life, and for moments of shared brainlessness. For common laughs, and your shoulders when eyes were filled with tears. Without you, my journey would not have been as joyful and meaningful as it was with your company. Thank you!

To **Jacintha Town**, for being a real friend to a complete stranger, me. For making me know that there is one more person in the hospital I might turn to in the times of hardships. Thank you!

To my most significant discovery, **Jad Assaad**. 'When you danced, the walls leaned closer to get closer to you*'. Shining diamond in the golden coat. Thank you for standing so firmly by my side through it all, like a cedar unshakable and strong. For your deep Presence, your pure Heart and engaging Mind, embracing reality with so much Passion, Grace, and Love like no one else does. For changing the rhythms, introducing the beats and sounds I had not heard before. Like the Full Moon, enhancing to be me. Thank you for being my Eyes, my Smile, my Heart.

Big thank you to all short but meaningful encounters during my last five years. Many other surnames shall remain in my heart. Some helped to shift the angle of my perspective, others introduced pure joy, or reminded about the reasons I work hard. Thank you for sharing all your beauty with me.

---

* *Dreaming boy* by Sarah Kay

REFERENCES

References

1. Sengpiel V, Elind E, Bacelis J, Nilsson S, Grove J, Myhre R, et al. Maternal caffeine intake during pregnancy is associated with birth weight but not with gestational length: results from a large prospective observational cohort study. BMC Med. 2013 Feb 19;11:42.
2. Barman M, Brantsæter AL, Nilsson S, Haugen M, Lundh T, Combs GF, et al. Maternal dietary selenium intake is associated with increased gestational length and decreased risk of preterm delivery. Br J Nutr. 2020 Jan 28;123(2):209–19.
3. Solé-Navais P, Brantsæter AL, Caspersen IH, Lundh T, Muglia LJ, Meltzer HM, et al. Maternal Dietary Selenium Intake during Pregnancy Is Associated with Higher Birth Weight and Lower Risk of Small for Gestational Age Births in the Norwegian Mother, Father and Child Cohort Study. Nutrients. 2020 Dec 23;13(1).
4. Maher B. Personal genomes: The case of the missing heritability. Nature. 2008 Nov 1;456(7218):18–21.
5. Hernández-Díaz S, Schisterman EF, Hernán MA. The Birth Weight "Paradox" Uncovered? Am J Epidemiol. 2006 Dec 1;164(11):1115–20.
6. Langley-Evans SC. Nutritional programming of disease: unravelling the mechanism. J Anat. 2009 Jul;215(1):36–51.
7. Dobzhansky T. Evolution, Genetics and Man. Second Printing edition. John Wiley & Sons Inc; 1957. 398 p.
8. Roll-Hansen N. Commentary: Wilhelm Johannsen and the problem of heredity at the turn of the 19th century. Int J Epidemiol. 2014 Aug;43(4):1007–13.
9. Jackson M, Marks L, May GHW, Wilson JB. The genetic basis of disease. Essays Biochem. 2018 Dec 3;62(5):643–723.
10. Falconer DS, Mackay TFC. Introduction to quantitative genetics. 4th edition. Pearson, Prentice Hall; 1996.
11. Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. Nat Rev Genet. 2013 Feb;14(2):139–49.
12. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era — concepts and misconceptions. Nature Reviews Genetics. 2008 Apr;9(4):255–66.
13. Sullivan PF. The Genetics of Schizophrenia. PLOS Medicine. 2005 Jul 26;2(7):e212.
14. Freitag CM. The genetics of autistic disorders and its clinical relevance: a review of the literature. Mol Psychiatry. 2007 Jan;12(1):2–22.
15. Meltzer HM, Brantsaeter AL, Ydersbond TA, Alexander J, Haugen M. Methodological challenges when monitoring the diet of pregnant women in a large study: experiences from the Norwegian Mother and Child Cohort Study (MoBa). Matern Child Nutr. 2008 Jan;4(1):14–27.
16. Darby WJ, Densen PM, Cannon RO, Bridgfort E, Martin MP, Kaser MM, et al. The Vanderbilt cooperative study of maternal and infant nutrition. I. Background. II. Methods. III. Description of the sample and data. J Nutr. 1953 Dec 10;51(4):539–63.

17. McGANITY WJ, Cannon RO, Bridgforth EB, Martin MP, Densen PM, Newbill JA, et al. The Vanderbilt cooperative study of maternal and infant nutrition. V. Description and outcome of obstetric sample. Am J Obstet Gynecol. 1954 Mar;67(3):491–500.
18. Thomson AM. Diet in pregnancy. I. Dietary survey technique and the nutritive value of diets taken by primigravidae. Br J Nutr. 1958;12(4):446–61.
19. Thomson AM. Diet in pregnancy. 3. Diet in relation to the course and outcome of pregnancy. Br J Nutr. 1959;13:509–25.
20. Susser M, Stein Z. Timing in prenatal nutrition: a reprise of the Dutch Famine Study. Nutr Rev. 1994 Mar;52(3):84–94.
21. Smith CA. The effect of wartime starvation in Holland upon pregnancy and its product. Am J Obstet Gynecol. 1947 Apr;53(4):599–608.
22. Barker DJ, Hales CN, Fall CH, Osmond C, Phipps K, Clark PM. Type 2 (non-insulin-dependent) diabetes mellitus, hypertension and hyperlipidaemia (syndrome X): relation to reduced fetal growth. Diabetologia. 1993 Jan;36(1):62–7.
23. Hales CN, Barker DJ, Clark PM, Cox LJ, Fall C, Osmond C, et al. Fetal and infant growth and impaired glucose tolerance at age 64. BMJ. 1991 Oct 26;303(6809):1019–22.
24. Roseboom TJ, van der Meulen JH, Ravelli AC, Osmond C, Barker DJ, Bleker OP. Effects of prenatal exposure to the Dutch famine on adult disease in later life: an overview. Mol Cell Endocrinol. 2001 Dec 20;185(1–2):93–8.
25. Vickers MH, Breier BH, McCarthy D, Gluckman PD. Sedentary behavior during postnatal life is determined by the prenatal environment and exacerbated by postnatal hypercaloric nutrition. Am J Physiol Regul Integr Comp Physiol. 2003 Jul;285(1):R271-273.
26. Gambling L, Dunford S, Wallace DI, Zuur G, Solanky N, Srai SKS, et al. Iron deficiency during pregnancy affects postnatal blood pressure in the rat. J Physiol. 2003 Oct 15;552(Pt 2):603–10.
27. Andersen HS, Gambling L, Holtrop G, McArdle HJ. Maternal iron deficiency identifies critical windows for growth and cardiovascular development in the rat postimplantation embryo. J Nutr. 2006 May;136(5):1171–7.
28. Khan IY, Dekou V, Douglas G, Jensen R, Hanson MA, Poston L, et al. A high-fat diet during rat pregnancy or suckling induces cardiovascular dysfunction in adult offspring. Am J Physiol Regul Integr Comp Physiol. 2005 Jan;288(1):R127-133.
29. Daenzer M, Ortmann S, Klaus S, Metges CC. Prenatal high protein exposure decreases energy expenditure and increases adiposity in young rats. J Nutr. 2002 Feb;132(2):142–4.
30. Pijnenborg R. Trophoblast invasion. Reproductive Medicine Review. 1994 Mar;3(1):53–73.
31. McCusker RR, Goldberger BA, Cone EJ. Caffeine content of specialty coffees. J Anal Toxicol. 2003 Oct;27(7):520–2.
32. Sökmen B, Armstrong LE, Kraemer WJ, Casa DJ, Dias JC, Judelson DA, et al. Caffeine Use in Sports: Considerations for the Athlete. The Journal of Strength & Conditioning Research. 2008 May;22(3):978–86.

33. Ullrich S, de Vries YC, Kühn S, Repantis D, Dresler M, Ohla K. Feeling smart: Effects of caffeine and glucose on cognition, mood and self-judgment. Physiol Behav. 2015 Nov 1;151:629–37.
34. Arendash GW, Schleif W, Rezai-Zadeh K, Jackson EK, Zacharia LC, Cracchiolo JR, et al. Caffeine protects Alzheimer's mice against cognitive impairment and reduces brain beta-amyloid production. Neuroscience. 2006 Nov 3;142(4):941–52.
35. Desbrow B, Leveritt M. Awareness and use of caffeine by athletes competing at the 2005 Ironman Triathlon World Championships. Int J Sport Nutr Exerc Metab. 2006 Oct;16(5):545–58.
36. Trexler ET, Smith-Ryan AE. Creatine and Caffeine: Considerations for Concurrent Supplementation. Int J Sport Nutr Exerc Metab. 2015 Dec;25(6):607–23.
37. Goldstein ER, Ziegenfuss T, Kalman D, Kreider R, Campbell B, Wilborn C, et al. International society of sports nutrition position stand: caffeine and performance. Journal of the International Society of Sports Nutrition. 2010 Jan 27;7(1):5.
38. Nawrot P, Jordan S, Eastwood J, Rotstein J, Hugenholtz A, Feeley M. Effects of caffeine on human health. Food Addit Contam. 2003 Jan;20(1):1–30.
39. Kolahdouzan M, Hamadeh MJ. The neuroprotective effects of caffeine in neurodegenerative diseases. CNS Neurosci Ther. 2017 Apr;23(4):272–90.
40. Poole R, Kennedy OJ, Roderick P, Fallowfield JA, Hayes PC, Parkes J. Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes. BMJ. 2017 Nov 22;359:j5024.
41. Qian J, Chen Q, Ward SM, Duan E, Zhang Y. Impacts of Caffeine during Pregnancy. Trends Endocrinol Metab. 2020 Mar;31(3):218–27.
42. Weng X, Odouli R, Li D-K. Maternal caffeine consumption during pregnancy and the risk of miscarriage: a prospective cohort study. Am J Obstet Gynecol. 2008 Mar;198(3):279.e1-8.
43. Group BMJP. Maternal caffeine intake during pregnancy and risk of fetal growth restriction: a large prospective observational study. BMJ. 2008 Nov 3;337:a2332.
44. Chen L-W, Wu Y, Neelakantan N, Chong MF-F, Pan A, van Dam RM. Maternal caffeine intake during pregnancy is associated with risk of low birth weight: a systematic review and dose-response meta-analysis. BMC Med. 2014 Sep 19;12:174.
45. Chen L-W, Fitzgerald R, Murrin CM, Mehegan J, Kelleher CC, Phillips CM, et al. Associations of maternal caffeine intake with birth outcomes: results from the Lifeways Cross Generation Cohort Study. Am J Clin Nutr. 2018 Dec 1;108(6):1301–8.
46. Galéra C, Bernard JY, van der Waerden J, Bouvard M-P, Lioret S, Forhan A, et al. Prenatal Caffeine Exposure and Child IQ at Age 5.5 Years: The EDEN Mother-Child Cohort. Biol Psychiatry. 2016 Nov 1;80(9):720–6.
47. Li D-K, Ferber JR, Odouli R. Maternal caffeine intake during pregnancy and risk of obesity in offspring: a prospective cohort study. Int J Obes (Lond). 2015 Apr;39(4):658–64.

48. Bajaj M, Eiche E, Neumann T, Winter J, Gallert C. Hazardous concentrations of selenium in soil and groundwater in North-West India. Journal of hazardous materials. 2011 May 1;189:640–6.
49. US EPA 1980 in Ambient Water Quality Criteria for Selenium (US EPA, Cincinnati, OH).
50. Combs GF. Selenium in global food systems. Br J Nutr. 2001 May;85(5):517–47.
51. Ngo DB, Dikassa L, Okitolonda W, Kashala TD, Gervy C, Dumont J, et al. Selenium status in pregnant women of a rural population (Zaire) in relationship to iodine deficiency. Trop Med Int Health. 1997 Jun;2(6):572–81.
52. Nordic Nutrition Recommendations 2012. :629.
53. Institute of Medicine (US) Subcommittee on Interpretation and Uses of Dietary Reference Intakes, Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. DRI Dietary Reference Intakes: Applications in Dietary Assessment
54. Kipp AP, Strohm D, Brigelius-Flohé R, Schomburg L, Bechthold A, Leschik-Bonnet E, et al. Revised reference values for selenium intake. Journal of Trace Elements in Medicine and Biology. 2015 Oct 1;32:195–9.
55. Nutrient Reference Values for Australia and New Zealand Including Recommended Dietary Intakes | NHMRC
56. SACN Statement on Selenium and Health - 2013
57. Sasaki S. Dietary Reference Intakes (DRIs) in Japan.
58. Spiller HA, Pfiefer E. Two fatal cases of selenium toxicity. Forensic Sci Int. 2007 Aug 24;171(1):67–72.
59. Hathcock JN. Vitamins and minerals: efficacy and safety. Am J Clin Nutr. 1997 Aug;66(2):427–37.
60. Nuttall KL. Evaluating selenium poisoning. Ann Clin Lab Sci. 2006;36(4):409–20.
61. Thomson CD. Selenium. In: Caballero B, editor. Encyclopedia of Human Nutrition (Third Edition)
62. Wasowicz W, Wolkanin P, Bednarski M, Gromadzinska J, Sklodowska M, Grzybowska K. Plasma trace element (Se, Zn, Cu) concentrations in maternal and umbilical cord blood in Poland. Relation with birth weight, gestational age, and parity. Biol Trace Elem Res. 1993 Aug;38(2):205–15.
63. Zachara BA, Wąsowicz W, Gromadzińska J, Skłodowska M, Krasomski G. Glutathione peroxidase activity, selenium, and lipid peroxide concentrations in blood from a healthy Polish population : I. Maternal and cord blood. Biol Trace Elem Res. 1986 Sep;10(3):175–87.
64. Pieczyńska J, Grajeta H. The role of selenium in human conception and pregnancy. J Trace Elem Med Biol. 2015 Jan;29:31–8.
65. Kullinger M, Granfors M, Kieler H, Skalkidou A. Adherence to Swedish national pregnancy dating guidelines and management of discrepancies between pregnancy dating methods: a survey study. Reprod Health
66. Committee Opinion No 700: Methods for Estimating the Due Date. Obstetrics & Gynecology. 2017 May;129(5):e150.
67. Stephansson O. 2015 annual report. Graviditetsregistret;

68. ACOG Committee Opinion No 579: Definition of term pregnancy. Obstet Gynecol. 2013 Nov;122(5):1139–40.
69. Kramer MS, McLean FH, Boyd ME, Usher RH. The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. JAMA. 1988 Dec 9;260(22):3306–8.
70. Bergsjø P, Denman DW, Hoffman HJ, Meirik O. Duration Of Human Singleton Pregnancy: A Population-based Study. Acta Obstetricia et Gynecologica Scandinavica. 1990 Jan 1;69(3):197–207.
71. Fleischman AR, Oinuma M, Clark SL. Rethinking the definition of "term pregnancy." Obstet Gynecol. 2010 Jul;116(1):136–9.
72. Quinn J-A, Munoz FM, Gonik B, Frau L, Cutland C, Mallett-Moore T, et al. Preterm birth: Case definition & guidelines for data collection, analysis, and presentation of immunisation safety data. Vaccine. 2016 Dec 1;34(49):6047–56.
73. Kramer MS, Demissie K, Yang H, Platt RW, Sauvé R, Liston R. The contribution of mild and moderate preterm birth to infant mortality. Fetal and Infant Health Study Group of the Canadian Perinatal Surveillance System. JAMA. 2000 Aug 16;284(7):843–9.
74. Goldenberg RL, Jobe AH. Prospects for research in reproductive health and birth outcomes. JAMA. 2001 Feb 7;285(5):633–9.
75. Ananth CV, Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. J Matern Fetal Neonatal Med. 2006 Dec;19(12):773–82.
76. Aylward GP. Neurodevelopmental outcomes of infants born prematurely. J Dev Behav Pediatr. 2005 Dec;26(6):427–40.
77. Costeloe K, EPICure Study Group. EPICure: facts and figures: why preterm labour should be treated. BJOG. 2006 Dec;113 Suppl 3:10–2.
78. Marlow N, Wolke D, Bracewell MA, Samara M, EPICure Study Group. Neurologic and developmental disability at six years of age after extremely preterm birth. N Engl J Med. 2005 Jan 6;352(1):9–19.
79. Siltanen M, Savilahti E, Pohjavuori M, Kajosaari M. Respiratory symptoms and lung function in relation to atopy in children born preterm. Pediatr Pulmonol. 2004 Jan;37(1):43–9.
80. Islam JY, Keller RL, Aschner JL, Hartert TV, Moore PE. Understanding the Short- and Long-Term Respiratory Outcomes of Prematurity and Bronchopulmonary Dysplasia. Am J Respir Crit Care Med. 2015 Jul 15;192(2):134–56.
81. Doyle LW, Cheung MM, Ford GW, Olinsky A, Davis NM, Callanan C. Birth weight <1501 g and respiratory health at age 14. Arch Dis Child. 2001 Jan;84(1):40–4.
82. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. Am J Respir Crit Care Med. 2001 Jun;163(7):1723–9.
83. Hård AL, Niklasson A, Svensson E, Hellström A. Visual function in school-aged children born before 29 weeks of gestation: a population-based study. Dev Med Child Neurol. 2000 Feb;42(2):100–5.
84. Jacobson LK, Dutton GN. Periventricular leukomalacia: an important cause of visual and ocular motility dysfunction in children. Surv Ophthalmol. 2000 Aug;45(1):1–13.

85. Olesen AW, Westergaard JG, Olsen J. Perinatal and maternal complications related to postterm delivery: a national register-based study, 1978-1993. Am J Obstet Gynecol. 2003 Jul;189(1):222–7.
86. Roos N, Sahlin L, Ekman-Ordeberg G, Kieler H, Stephansson O. Maternal risk factors for postterm pregnancy and cesarean delivery following labor induction. Acta Obstet Gynecol Scand. 2010 Aug;89(8):1003–10.
87. van Zijl MD, Koullali B, Mol BW, Pajkrt E, Oudijk MA. Prevention of preterm delivery: current challenges and future prospects. Int J Womens Health. 2016 Oct 31;8:633–45.
88. Biggio JR. Current Approaches to Risk Assessment and Prevention of Preterm Birth—A Continuing Public Health Crisis. Ochsner J. 2020;20(4):426–33.
89. Socialstyrelsen. (2015). Graviditeter, förlossningar och nyfödda barn, Medicinska födelseregistret 1973–2014, Assisterad befruktning 1991– 2013.
90. Pickett KE, Abrams B, Selvin S. Defining preterm delivery--the epidemiology of clinical presentation. Paediatr Perinat Epidemiol. 2000 Oct;14(4):305–8.
91. Berkowitz GS, Blackmore-Prince C, Lapinski RH, Savitz DA. Risk factors for preterm birth subtypes. Epidemiology. 1998 May;9(3):279–85.
92. Lumley J. The epidemiology of preterm birth. Baillieres Clin Obstet Gynaecol. 1993 Sep;7(3):477–98.
93. Goldenberg. Epidemiology and causes of preterm birth
94. Thompson JMD, Irgens LM, Rasmussen S, Daltveit AK. Secular trends in socio-economic status and the implications for preterm birth. Paediatr Perinat Epidemiol. 2006 May;20(3):182–7.
95. Peacock JL, Bland JM, Anderson HR. Preterm delivery: effects of socioeconomic factors, psychological stress, smoking, alcohol, and caffeine. BMJ. 1995 Aug 26;311(7004):531–5.
96. Parker JD, Schoendorf KC, Kiely JL. Associations between measures of socioeconomic status and low birth weight, small for gestational age, and premature delivery in the United States. Ann Epidemiol. 1994 Jul;4(4):271–8.
97. Basso O, Wilcox A. Mortality risk among preterm babies: Immaturity vs. underlying pathology. Epidemiology. 2010 Jul;21(4):521–7.
98. Pickett KE, Ahern JE, Selvin S, Abrams B. Neighborhood socioeconomic status, maternal race and preterm delivery: a case-control study. Ann Epidemiol. 2002 Aug;12(6):410–8.
99. Lobel M, Dunkel-Schetter C, Scrimshaw SC. Prenatal maternal stress and prematurity: a prospective study of socioeconomically disadvantaged women. Health Psychol. 1992;11(1):32–40.
100. Smith GCS, Pell JP, Dobbie R. Interpregnancy interval and risk of preterm birth and neonatal death: retrospective cohort study. BMJ. 2003 Aug 9;327(7410):313.
101. DeFranco EA, Stamilio DM, Boslaugh SE, Gross GA, Muglia LJ. A short interpregnancy interval is a risk factor for preterm birth and its recurrence. American Journal of Obstetrics and Gynecology. 2007 Sep 1;197(3):264.e1-264.e6.
102. Conde-Agudelo A, Rosas-Bermúdez A, Kafury-Goeta AC. Birth spacing and risk of adverse perinatal outcomes: a meta-analysis. JAMA. 2006 Apr 19;295(15):1809–23.

103. Cobo T, Kacerovsky M, Jacobsson B. Risk factors for spontaneous preterm delivery. International Journal of Gynecology & Obstetrics. 2020;150(1):17–23.
104. Rodrigues J, Niederman MS. Pneumonia complicating pregnancy. Clin Chest Med. 1992 Dec;13(4):679–91.
105. Moore VM, Davies MJ. Diet during pregnancy, neonatal outcomes and later health. Reprod Fertil Dev. 2005;17(3):341–8.
106. Agrawal V, Hirsch E. Intrauterine infection and preterm labor. Semin Fetal Neonatal Med. 2012 Feb;17(1):12–9.
107. Iams JD, Goldenberg RL, Mercer BM, Moawad AH, Meis PJ, Das AF, et al. The preterm prediction study: can low-risk women destined for spontaneous preterm birth be identified? Am J Obstet Gynecol. 2001 Mar;184(4):652–5.
108. Catley C, Frize M, Walker CR, Petriu DC. Predicting high-risk preterm birth using artificial neural networks. IEEE Trans Inf Technol Biomed. 2006 Jul;10(3):540–9.
109. Magnus P, Bakketeig LS, Skjaerven R. Correlations of birth weight and gestational age across generations. Ann Hum Biol. 1993 Jun;20(3):231–8.
110. Klebanoff MA, Meirik O, Berendes HW. Second-generation consequences of small-for-dates birth. Pediatrics. 1989 Aug;84(2):343–7.
111. Porter TF, Fraser AM, Hunter CY, Ward RH, Varner MW. The risk of preterm birth across generations. Obstet Gynecol. 1997 Jul;90(1):63–7.
112. Winkvist A. Familial patterns in birth characteristics: impact on individual and population risks. International Journal of Epidemiology. 1998 Apr 1;27(2):248–54.
113. Boyd HA, Poulsen G, Wohlfahrt J, Murray JC, Feenstra B, Melbye M. Maternal Contributions to Preterm Delivery. Am J Epidemiol. 2009 Dec 1;170(11):1358–64.
114. Cutland CL, Lackritz EM, Mallett-Moore T, Bardají A, Chandrasekaran R, Lahariya C, et al. Low birth weight: Case definition & guidelines for data collection, analysis, and presentation of maternal immunization safety data. Vaccine. 2017 Dec 4;35(48 Pt A):6492–500.
115. Wilcox AJ. On the importance--and the unimportance--of birthweight. Int J Epidemiol. 2001 Dec;30(6):1233–41.
116. Skjærven R, Gjessing HK, Bakketeig LS. Birthweight by gestational age in Norway. Acta Obstetricia et Gynecologica Scandinavica. 2000;79(6):440–9.
117. Gardosi J. New definition of small for gestational age based on fetal growth potential. Horm Res. 2006;65 Suppl 3:15–8.
118. Kaufmann P, Black S, Huppertz B. Endovascular trophoblast invasion: implications for the pathogenesis of intrauterine growth retardation and preeclampsia. Biol Reprod. 2003 Jul;69(1):1–7.
119. Mayer C, Joseph KS. Fetal growth: a review of terms, concepts and issues relevant to obstetrics. Ultrasound Obstet Gynecol. 2013 Feb;41(2):136–45.
120. Marsál K, Persson PH, Larsen T, Lilja H, Selbing A, Sultan B. Intrauterine growth curves based on ultrasonically estimated foetal weights. Acta Paediatr. 1996 Jul;85(7):843–8.
121. Gardosi J. Customized growth curves. Clin Obstet Gynecol. 1997 Dec;40(4):715–22.

122. Ekbom A. The Swedish Multi-generation Register. In: Dillner J, editor. Methods in Biobanking
123. The Swedish Medical Birth Register. Socialstyrelsen.
124. Källén B, Källén K. The Swedish Medical Birth Register - A Summary of Content and Quality
125. Nilsen RM, Vollset SE, Gjessing HK, Skjaerven R, Melve KK, Schreuder P, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. Paediatr Perinat Epidemiol. 2009 Nov;23(6):597–608.
126. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. Acta Obstetricia et Gynecologica Scandinavica. 2000;79(6):435–9.
127. Ridings JE. The thalidomide disaster, lessons from the past. Methods Mol Biol. 2013;947:575–86.
128. Vargesson N. Thalidomide-induced teratogenesis: History and mechanisms. Birth Defects Res C Embryo Today. 2015 Jun;105(2):140–56.
129. Irgens LM. The Medical Birth Registry of Norway; a source for epidemiological and clinical research. Scand J Rheumatol Suppl. 1998;107:105–8.
130. Wu W, Witherspoon DJ, Fraser A, Clark EAS, Rogers A, Stoddard GJ, et al. The heritability of gestational age in a two-million member cohort: implications for spontaneous preterm birth. Hum Genet. 2015 Jul;134(7):803–8.
131. Clausson B, Lichtenstein P, Cnattingius S. Genetic influence on birthweight and gestational length determined by studies in offspring of twins. BJOG: An International Journal of Obstetrics & Gynaecology. 2000;107(3):375–81.
132. Treloar SA, Macones GA, Mitchell LE, Martin NG. Genetic influences on premature parturition in an Australian twin sample. Twin Res. 2000 Jun;3(2):80–2.
133. Lunde A, Melve KK, Gjessing HK, Skjaerven R, Irgens LM. Genetic and Environmental Influences on Birth Weight, Birth Length, Head Circumference, and Gestational Age by Use of Population-based Parent-Offspring Data. American Journal of Epidemiology. 2007 Feb 19;165(7):734–41.
134. Kistka ZA-F, DeFranco EA, Ligthart L, Willemsen G, Plunkett J, Muglia LJ, et al. Heritability of parturition timing: an extended twin design analysis. American Journal of Obstetrics and Gynecology. 2008 Jul;199(1):43.e1-43.e5.
135. Svensson AC, Sandin S, Cnattingius S, Reilly M, Pawitan Y, Hultman CM, et al. Maternal effects for preterm birth: a genetic epidemiologic study of 630,000 families. Am J Epidemiol. 2009 Dec 1;170(11):1365–72.
136. York TP, Eaves LJ, Lichtenstein P, Neale MC, Svensson A, Latendresse S, et al. Fetal and Maternal Genes' Influence on Gestational Age in a Quantitative Genetic Analysis of 244,000 Swedish Births. Am J Epidemiol. 2013 Aug 15;178(4):543–50.
137. York TP, Strauss JF, Eaves LJ. A narrow heritability evaluation of gestational age at birth. Hum Genet. 2015 Jul;134(7):809–11.
138. Maurel M-C, Kanellopoulos-Langevin C. Heredity--venturing beyond genetics. Biol Reprod. 2008 Jul;79(1):2–8.
139. Hollingsworth K, Bowyer KW, Lagree S, Fenker SP, Flynn PJ. Genetically identical irises have texture similarity that is not detected by iris biometrics. Computer Vision and Image Understanding. 2011 Nov;115(11):1493–502.

140. Lewontin RC. The analysis of variance and the analysis of causes. International Journal of Epidemiology. 2006 Jun 1;35(3):520–5.
141. Anderson TM, Lavista Ferres JM, Ren SY, Moon RY, Goldstein RD, Ramirez J-M, et al. Maternal Smoking Before and During Pregnancy and the Risk of Sudden Unexpected Infant Death. Pediatrics
142. Robinson LD, Jewell NP. Some Surprising Results about Covariate Adjustment in Logistic Regression Models. International Statistical Review / Revue Internationale de Statistique. 1991;59(2):227–40.
143. Allison P. Comparing Logit and Probit Coefficients Across Groups. Sociological Methods & Research - SOCIOL METHOD RES. 1999 Nov 1;28:186–208.