From the Department of Medical Epidemiology and Biostatistics
Karolinska Institutet, Stockholm, Sweden

REGISTER-BASED STUDIES OF INFERTILITY AND THE USE OF ASSISTED REPRODUCTIVE TECHNIQUES — POSSIBILITIES AND CHALLENGES

Chen Wang

王晨

Stockholm 2022

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2022

ISBN 978-91-8016-728-4
Cover illustration: Gouache painting ‘Newborn’ by Chen Wang

Register-based studies of infertility and the use of assisted reproductive techniques — possibilities and challenges
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Chen Wang

The thesis will be defended in public at lecture hall Atrium, Nobels väg 12 B, Karolinska Institutet, Solna, August 26th, 2022, 1:00 PM

Principal Supervisor:
Docent Anna Sara Oberg
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics
Harvard T. H. Chan School of Public Health
Department of Epidemiology

Co-supervisor(s):
Professor Sonia Hernández-Díaz
Harvard T. H. Chan School of Public Health
Department of Epidemiology

Professor Catarina Almqvist Malmros
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics
Karolinska University Hospital
Astrid Lindgren Children's Hospital

Docent Anna L. V. Johansson
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Opponent:
Professor Anne-Marie Nyboe Andersen
University of Copenhagen
Department of Public Health

Examination Board:
Professor Anders Hjern
Karolinska Institutet
Department of Medicine, Solna
Division of Clinical Epidemiology

Docent Elizabeth Arkema
Karolinska Institutet
Department of Medicine, Solna
Division of Clinical Epidemiology

Docent Emir Henic
Lund University
Department of Translational Medicine
Division of Reproductive medicine, Malmö

To all those who love me or have helped me

致所有爱我和帮助过我的人

ABSTRACT

The implementation of assisted reproductive techniques (ART) has enabled over 10 million births in the past 45 years. Concern for excess risks of adverse birth outcomes has largely, but not entirely, been mitigated by practice recommendations for single embryo transfer. Evaluations of long-term health are still relatively limited but increasingly warranted as the number of children conceived with ART that have reached adolescence and adulthood is growing.

In this thesis, we used nationwide Swedish register data to assess the long-term health of children conceived with ART, with the specific aim of disentangling the potential role of the intervention from its indication, i.e., underlying infertility and associated parental background characteristics. Linkage of multiple registers from different sources enabled prospectively recorded follow-up of sufficient length, consideration of a wide range of parental background characteristics, and low risk of selection bias. Focusing on outcomes for which evidence was limited due to few or poorly performed studies, we investigated attention deficit hyperactivity disorder (ADHD) and poor school performance in Study I; depression, anxiety, antidepressant use, and suicidal behavior in Study II; and childhood asthma in Study III. In addition to examining how much of the potential observed differences in these outcomes could be explained by parental factors, we also explored the potential influence of specific procedures such as intra-cytoplasmic sperm injection (ICSI) and frozen embryo transfer.

In Study I, we found that children conceived with ART were at lower risk of ADHD and they performed better in grade 9 (higher average grade and eligibility for upper secondary school) compared with all other children. Adjustment for parental background characteristics reversed the findings such that children conceived with ART appeared at a slight disadvantage. However, when restricting the contrast to children of couples with known infertility, no differences were seen in any of the outcomes. In Study II, we applied a similar approach to evaluate mental health outcomes, extending the follow-up period into early adulthood. Adolescents conceived with ART were not at elevated risk of depression or suicidal behavior compared to other adolescents (irrespective of parental infertility). Compared with all others, adolescents conceived with ART had a modestly elevated risk of obsessive-compulsive disorder, but this was greatly attenuated by adjustment for parental factors and not seen at all in comparison to adolescents born to couples with known infertility. In Study III, the previously reported link between ART and childhood asthma was put to even further scrutiny by including also a comparison of first-born maternal cousins. Once again, an observed small excess risk appeared to be largely explained by parental background factors. However, a modestly elevated risk of asthma in the first year of life seen only in children born after frozen embryo transfer warrants further confirmation.

In Study IV, we shifted focus from casual investigation of children's health to prediction of pregnancy complications in women undergoing ART, developing pre- and post-treatment prediction models for preeclampsia, placental complications (accrete, previa, and abruption), and postpartum hemorrhage. A wide range of predictors was considered, including demographic factors, prior medical history and medication use, and ART treatment cycle information. Several different model algorithms were implemented and cross-compared, including logistic regression, random forest, and gradient boosting. Still, the final model prediction performances were just above 60% at best (ranging from 54% to 62%). Across the different complications, the most influential predictors tended to be body mass index (BMI), parental age, treatment year, region of residence, infertility diagnosis, and type of embryo transfer (in additional attempts). While some gain may be made from further refinements of the predictor categorization, the most critical improvement for continued prediction efforts should entail expansion to consider also repeated clinical measurements in pregnancy.

In conclusion, these studies provide overall reassurance with regard to childhood asthma and the long-term neurodevelopmental and mental health of individuals conceived with ART. Some excess risks seen in relation to the general population appear to be largely driven by differences in parental characteristics, stressing the importance of considering the role of the underlying indication (and its associated risk factors) in ART safety studies. Further studies are required to combine treatment-related information with clinical measurements throughout pregnancy to improve the prediction of maternal complications and tailor a clinical protocol for this specific population at risk.

LIST OF SCIENTIFIC PAPERS

I. Wang C, Johansson A, Rodriguez-Wallberg K, Almqvist C, Hernández-Díaz S, Öberg S. (2021). **Assisted Reproductive Techniques, ADHD, and School Performance.** *Pediatrics*, 148(1). https://doi.org/10.1542/peds.2020-033183

II. Wang C, Johansson A, Rodriguez-Wallberg K, Landén M, Almqvist C, Hernández-Díaz S, Öberg S. (2022). **Long-term Follow-up of Psychiatric Disorders in Children and Adolescents Conceived by Assisted Reproductive Techniques in Sweden.** *JAMA Psychiatry*, 79(2), 133. https://doi.org/10.1001/jamapsychiatry.2021.3647

III. Wang C, Johansson A, Rodriguez-Wallberg K, Hernández-Díaz S, Almqvist C, Öberg S. **Are assisted reproductive techniques responsible for an elevated risk of asthma in children?** *Manuscript.*

IV. Wang C, Johansson A, Nyberg C, Pareek A, Almqvist C, Hernández-Díaz S, Öberg S. **Prediction models of pregnancy complications in women undergoing assisted reproduction, using machine learning methods.** *Manuscript.*

CONTENTS

LIST OF ABBREVIATIONS

ADHD	Attention Deficit Hyperactivity Disorder
ART	Assisted Reproductive Technique
ATC	Anatomical Therapeutical Chemical (classification system for pharmaceutical drugs)
ASD	Autism Spectrum Disorder
AUC	Area Under Curve
BMI	Body Mass Index
CDR	Cause of Death Register
CI	Confidence Interval
DAG	Directed Acyclic Graph
DT	Decision Tree
HR	Hazard Ratio
GB	Gradient Boosting
GDPR	General Data Protection Regulation
GLM	Generalized Linear Models
LISA	Longitudinal Integrated Database for Health Insurance and Labor Market Studies (longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknads-studier)
LR	Logistic Regression
ICD	International Classification of Diseases
ICSI	Intracytoplasmic Sperm Injection
IVF	In Vitro Fertilization
IUGR	Intrauterine Growth Restriction
IUI	Intrauterine Insemination
MAR	Missing At Random
MBR	Medical Birth Register
MCAR	Missing Completely At Random
MGR	Multi-Generation Register
MNAR	Missing Not At Random
NBC	Naïve Bayes Classification

NPR	National Patient Register
NSR	National School Register
OCD	Obsessive-Compulsive Disorder
OR	Odds Ratio
PCOS	Polycystic Ovary Syndrome
PDR	Prescribed Drug Register
PGD	Preimplantation Genetic Diagnosis
PPH	Postpartum Hemorrhage
Q-IVF	National Quality Register for Assisted Reproductive Techniques
RF	Random Forest
ROC	Receiver Operating Characteristics
RTB	Total Population Register (registret över totalbefolkningen)
SCB	Statistics Sweden (Statistiska Centralbyrån)
SDP	Smoking During Pregnancy
SHAP	Shapley Additive Explanations
SOS	National Board of Health and Welfare (Socialstyrelsen)
SVM	Support Vector Machine
WHO	World Health Organization

1 INTRODUCTION

1.1 INFERTILITY

The World Health Organization (WHO) characterizes clinical infertility as "a disease of the reproductive system defined by the failure to achieve a clinical pregnancy after 12 months or more of regular unprotected sexual intercourse".[1] Among young couples the probability of achieving pregnancy is around 20 to 25% per reproductive cycle,[2] and for couples without trouble conceiving the cumulative conception rate in a year is nearly 100%. Notably, the clinical definition of infertility concerns a couples' reproductive efficiency rather than their absolute ability, such that several couples with infertility may conceive without assistance.

Infertility is common, with more than 10% of couples experiencing trouble conceiving worldwide.[3] In worldwide assessments between 2015 and 2016, 20-30% of couples who have known reasons for infertility tend to be attributed to female factors, another 20-30% to male factors, and the remaining attributed to both.[4,5] Female disorders include ovulatory dysfunction, tubal or uterine disease, endometriosis, and chromosomal abnormalities, whereas the primary male factors concern impaired sperm quantity and quality.[6-8] Besides different medical conditions, several lifestyle factors, e.g., very low or high body mass index (BMI) and smoking, can lead to infertility. In women in particular, age is a critical determinant of the ability to conceive. As women age, the ovarian reserve naturally declines in both quantity (i.e. diminishing follicles in the ovary) and quality (i.e., oocyte quality).[9,10] In addition to a decreasing ability to conceive, the rate of miscarriage increases with advancing age.[11] The mean age at childbearing age has been increasing over the past decades,[12,13] reflecting in part a growing emphasis on educational as well as career goals.[14]

1.2 ASSISTED REPRODUCTIVE TECHNIQUES (ART)

Before in vitro fertilization (IVF), a few approaches were available to overcome infertility caused by specific factors. For instance, ovulation induction with clomiphene citrate or gonadotropin was developed to help restore ovulation in women with anovulatory cycles, e.g., due to polycystic ovary syndrome (PCOS),[15,16] tubal or uterine anomalies could be targeted with microsurgery, and male factor with artificial insemination. With the exception of artificial insemination, however, these could only be applied to specific problems, until the focus of infertility specialists turned to the possible use of extracorporeal fertilization.[17]

In 1977, a 30-year-old nulliparous married woman with a history of 9 years of infertility successfully conceived through fertilization in vitro and reimplantation of the embryo by Steptoe and Edwards.[18] As a result, a healthy, normal-weight infant girl was safely delivered in London on July 25, 1978, [18] sparking wonders, but also concerns, worldwide.[19,20]

A standard assisted conception with IVF involves several steps. It typically starts with daily injections of gonadotropins to stimulate the ovaries to produce many follicles. After this

controlled ovarian hyperstimulation, oocytes are retrieved from the matured follicles through ultrasound-guided transvaginal methods and placed with motile sperms in a Petri dish for fertilization. After being subsequently cultured for 2 to 5 days, one or more embryos will be selected and transferred back to the uterus.

Since the 2000s, the term "assisted reproductive techniques" (ARTs) has been used to collectively describe "all treatments or procedures that include the in vitro handling of both oocytes and sperm or of embryos for the purpose of establishing a pregnancy".[1] In the four decades since the first successful IVF, the efficiency of the procedures has continually improved as a result of fine-tuning treatment protocols and culture mediums, and the development of additional procedures. For instance, instead of culturing eggs together with motile sperm, individually selected sperm can be injected into the egg with the use of a thin glass pipette, through a procedure called intra-cytoplasmic sperm injection (ICSI). Introduced in the early 90s, ICSI was initially provided to increase the fertilization rate of couples with severe male factor infertility.[21] Due to its success, it has been increasingly used beyond its original intention to now account for nearly half of ART cycles.[22] Furthermore, the development of cryopreservation techniques has allowed excess embryos of good quality to be stored for subsequent implantation attempts without additional egg retrieval procedures. Initially, cleavage stage embryos were treated with cryoprotectant and cooled over a period of hours. This "slow-freezing" has gradually been replaced with "vitrification", in which a higher concentration of cryoprotectant is used to freeze embryos in a matter of seconds. The selected embryos are usually incubated to the blastocyst stage, as blastocysts are more robust and stable for both cryopreservation and embryo transfer. Today, more and more IVF clinics apply vitrification of blastocysts, which is generally considered to lead to better clinical results compared with conventional slow freezing.[23] Studies have also shown that vitrification provides very good results when applied to embryos in the cleavage stage.[24,25] In contrast to fresh embryos, which are transferred back in the same cycle that eggs were retrieved, frozen embryos are transferred at some later point. The options then are to synchronize the transfer with the natural ovulation cycle, help ovulation along with induction and/or trigger, or to instead inhibit the ovulation. Finally, early embryo biopsy and preimplantation genetic diagnosis (PGD) have been developed for couples predisposed to a serious genetic disease or chromosomal abnormality in order to minimize the risk of a fetal genetic defect and corresponding risk of having to undergo an abortion.

Today, there is no doubt that the treatment of infertility has been revolutionized by ART. The US national ART report in 2016 showed the average rate of clinical pregnancy for an IVF cycle with fresh, non-donor eggs to be 27.3%.[26] Due to the subsequent risk of miscarriage the birth rate drops to 22.2%, similar to the rate of natural conception in the general population.[26] The extensive implementation of ARTs has resulted in over 9 million births worldwide by 2019.[5] The groundbreaking intervention was also recognized with a Nobel prize awarded to Robert Edwards in 2010.

1.3 INFERTILITY EVALUATION AND ART TREATMENT IN SWEDEN

The goals of infertility evaluation are to set a diagnosis and provide a prognosis for the chance of pregnancy with and without treatment.[27] In general, heterosexual couples should have tried to achieve pregnancy through regular unprotected sexual intercourse for at least one year before the investigation. If there is a known or probable factor that reduces fertility (including advanced maternal age) the investigation can be carried out earlier. Since 2005, women in same-sex couples are eligible for evaluation and fertility assistance with donated sperm.[27]

A standard infertility evaluation includes assessments of the woman's ovulation history, ovarian reserve, uterine and tubal examination via vaginal ultrasound and the man's sperm sample, as well as the medical history and infectious disease screening of both. The potential influence of lifestyle factors such as over- or under-weight, tobacco smoking, alcohol and caffeine intake, psychological stress, and extreme physical activity should be considered early to allow the couple the opportunity to make changes.[27]

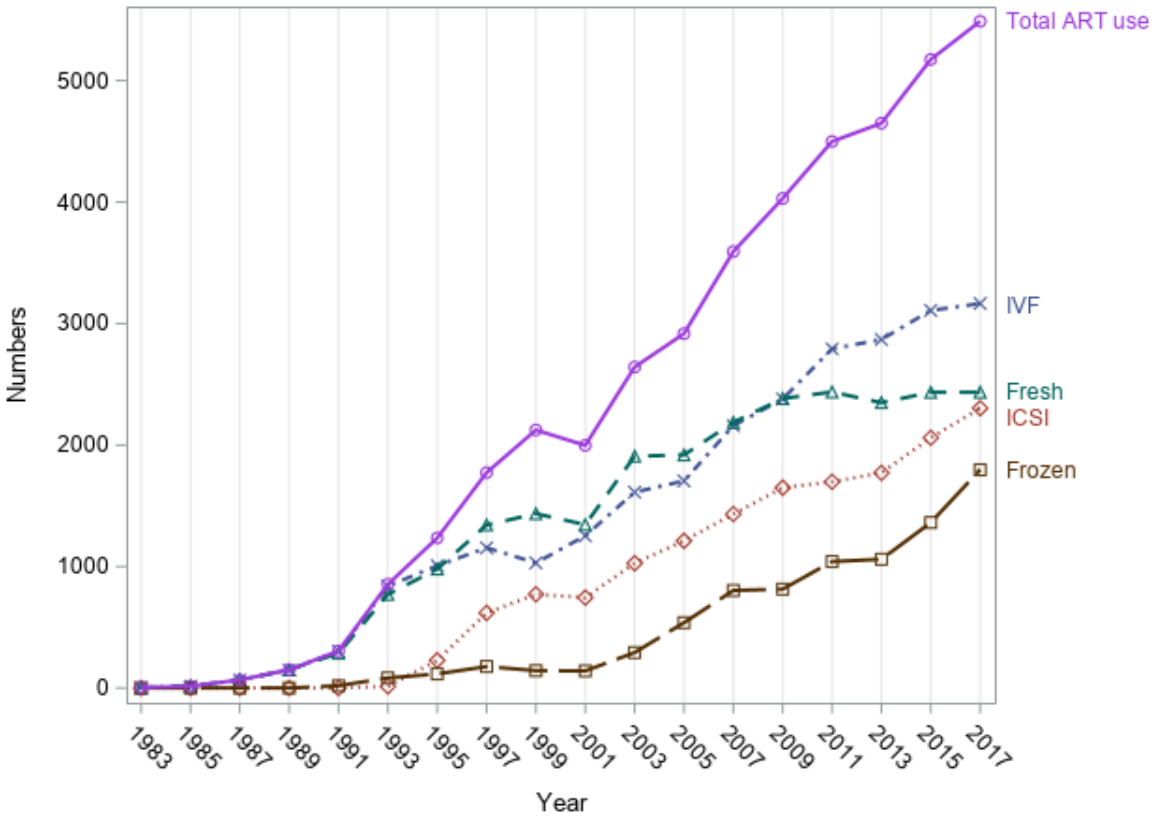
After the evaluation, individualized fertility treatment may be provided within the Swedish publicly funded healthcare system. Similar to the time before the introduction of ART, specific causes of infertility would first be addressed with microsurgery (tubal or uterine anomalies), ovulation induction (ovulatory dysfunction), or intrauterine insemination (unexplained or mild male factor infertility and sperm donations). If these methods fail, and for all those that do not qualify for them, ART should be recommended. In practice, due to the wide implementation and success of ART, it is often the first choice regardless. Although specific criteria can differ slightly between IVF clinics in Sweden, the basic requirements for publicly funded ART treatment often include: 1) no joint children; 2) at least 1 year of failing to conceive; 3) a certain range of women's BMI (normally between 18 and 35); 4) a certain age range, for women typically 25 to 40 and for men 25 to 56 years. The publicly funded ART treatment covers up to three ART attempts and the queue time varies from months to years depending on the clinic. The majority of IVF clinics in Sweden also offer privately funded ART services, for which some of the previously noted requirements can be relaxed.

Since the first birth following ART in Sweden in 1982, the following laws and policies have been introduced:[27]

- Sperm donation for insemination was regulated by law since 1985 and allowed in ART in 1988
- Egg donation was allowed in 2003
- Insemination and ART became available for lesbian couples in 2005
- Cryopreservation of embryos was available since the early 1990s. A maximum of 5 years of storage of frozen embryos was allowed in 1997 and recently extended to 10 years in 2019
- ART for single mothers was allowed in 2016
- Embryo donation regulation was introduced in 2019

Similar to the global implementation of ART, the number of births from ART in Sweden has increased dramatically since the 1990s (Figure 1.3). The widespread implementation and success have contributed to many couples discarding other treatment options on behalf of wishing to go directly to ART. Today, ARTs are responsible for more than 5000 births in Sweden each year (around 5.3%). Regarding the use of specific ART procedures, both ICSI and frozen embryo transfer started to be implemented in the early 1990s. In contrast to the immediate gradual increase in the use of ICSI, however, the use of frozen embryo transfer did not pick up until the early 2000s when a similar steady increase in the implementation took off. With the increasing reliance and success of frozen embryo transfers, the number of births from fresh embryo transfer has stagnated and remained relatively stable since the late 2000s.

Figure 1.3: Number of births from ART in Sweden from 1983 to 2017



Work of the figure was based on data from both the Swedish National Board of Health and Welfare and the Swedish National Quality Register for Assisted Reproductive Techniques (Q-IVF)

Although sperm donation has been allowed since the early days of ART and egg donation became legal in 2003, ART with donated gametes was only performed at university hospitals up to 2019, when also private clinics could offer such alternatives.[28] Births from donated gametes have thus far accounted for less than 5 percent of the total number of births from ART each year. In 2019, ART with donated sperm or eggs in Sweden resulted in 416 and 96 births respectively.[28] Another 328 births were the result of donor sperm insemination.

2 LITERATURE REVIEW

2.1 MATERNAL COMPLICATIONS DURING PREGNANCY AND DELIVERY

Preeclampsia is a serious pregnancy complication and the leading cause of perinatal and maternal morbidity and mortality.[29] Women with preeclampsia usually experience hypertension and proteinuria, which may lead to organ damage such as impaired kidney and liver function, and also increase the risk of brain injury and blood clotting problems. ART treatment has been associated with acute changes in maternal hemodynamic parameters including heart rate and cardiac function.[30] Consistent reports found women who conceive with ARTs at elevated risk of gestational hypertension and preeclampsia compared with women that conceived spontaneously.[31,32] Although the etiology remains unclear, it has been proposed that the elevated risk of preeclampsia could in part be due to the controlled ovarian hyperstimulation during ART treatment.[33] Conversely, the use of frozen embryo transfer has also been implicated, and in particular the ovulation inhibition protocol that precludes the development of a corpus luteum.[34-36] In contrast to such direct effects, it has been shown that much of the previously observed excess risk of preeclampsia following ART was mediated by the high occurrence of multiple gestations.[37]

Pregnancies following ART may also be at elevated risk of placental complications[38] and postpartum hemorrhage (PPH).[39] Placental complications such as placenta previa, placenta accreta, and placental abruption, are relatively rare but serious conditions that threaten the health of both mother and fetus. It has been proposed that a higher risk of placenta previa and placenta abruption in ART pregnancies is attributed to fresh embryo transfers (compared to frozen),[40] suggesting that the uterine environment at the time of embryo implantation is an important contributor in the development of these complications.[34,41] In contrast, several previous studies have reported an elevated risk of PPH following frozen embryo transfer, with putative mechanisms ranging from the aforementioned absence of corpus luteum to the excess risk of large-for-gestational-age delivery observed in this group.[34-36]

2.2 SHORT-TERM OFFSPRING OUTCOMES

2.2.1 Multiple gestation

After the first successful IVF birth, the transfer of more than one embryo was regarded as an essential method to increase the pregnancy rate. Not surprisingly, this introduced a dramatically high dizygotic twinning rate in pregnancies from IVF.[17] Evidence derived from a mathematical model [42] and a systematic review of four studies [43] have shown that women undergoing IVF had a 20- to 60-fold higher likelihood of multiple gestation under such practice. Although much more modest, studies have also pointed to a small elevated risk of monozygotic twinning following ART, especially in blastocyst embryo transfer (day 5; with incidence estimates ranging from 2- to 5-fold higher risk compared to spontaneous monozygotic twinning).[44,45] Multiple gestations only account for 1.5 % of all live births

worldwide. However, among IVF children in Europe 27.8% were twins in 1997,[46] and in the US around half (46%) were twins in 2001.[47]

Twin gestation is broadly appreciated as auspicious in some cultures, but it is also burdened with an increased risk of maternal and fetal complications. Carrying more than one fetus is associated with a greater need for bed rest, even though this appears to have little influence on established higher risks of preeclampsia, cesarean section, and PPH.[3] Twins and higher order multiples are at elevated risk of congenital malformations and are very likely to be born prematurely and small (also for their gestational age). Compared to singletons they also have a higher risk of dying in utero, at birth or in the perinatal period.

Findings from a systematic review of several RCTs have indicated that even though double embryo transfer results in a higher pregnancy rate per cycle (OR=1.94, 95% CI 1.47-2.55), a similar pregnancy rate can be obtained by single embryo transfer followed by a subsequent frozen-thawed embryo transfer (OR=1.19, 95% CI 0.87-1.62).[43] In 2001, a single-embryo transfer policy recommendation was introduced in Sweden and seemingly broadly implemented by 2003.[48] While generally adopted in Europe, the recommendation of single embryo transfer is not yet implemented worldwide.

2.2.2 Peri- and neonatal outcomes

Even among singletons, pregnancies from ART are at elevated risk of adverse pre- and perinatal outcomes. Two separate meta-analyses found that compared with spontaneously conceived singletons, IVF was associated with around 1.4- to 2-fold odds of each of the perinatal complications examined; preterm delivery, low birth weight, and small for gestational age.[49,50] Birth defects also appear to be more common in infants conceived with IVF, with reports of a 30-40% higher risk of major malformations compared with spontaneously conceived infants.[51-53] Possible explanations for observed excess risks include a more meticulous screening of babies conceived with IVF (ascertainment bias), influence of parental characteristics related to the underlying infertility (confounding)[51,54-56], or possible adverse effects of the interventions per se. The process of culturing and developing embryos might lead to epigenetic changes, thereby possibly impacting imprinting, whereby epigenetic reprogramming in the germ line results in gene expression according to parental origin. Children born from IVF have been found at higher risk of any imprinting disorder.[57] Intrinsic errors in imprinting have already been documented in oocytes from women undergoing ovulation induction [58] and in sperm from men with suboptimal fertility.[59] It is thus possible that the use of ART allows the errors that already exist in this population to be accentuated and/or carried forward.[60]

2.3 LONG-TERM OFFSPRING OUTCOMES

2.3.1 Neurodevelopmental outcomes

Infants and toddlers conceived with ART seem to display normal functioning with respect to neuromotor and cognitive skills, language, and behavior, according to the conclusions of several systematic reviews.[61-63] However, as also suggested in one of these reviews, many studies were small (fewer than 500 children) and outcomes were ascertained by self-report, interviews, and questionnaires to mothers or teachers.[62] Modest risks may thus have been missed, and findings influenced by information bias. Early evidence on autism spectrum disorders (ASD) was highly inconsistent due to reports of both positive and negative associations.[64,65] A later review suggested no difference in risk of ASD among ART toddlers, however, the conclusion was based on four out of seven studies.[66] The conclusion was subsequently confirmed by a large register-based study in Sweden, although this did note a small but statistically significant risk of intellectual disability.[67]

A few studies have been conducted with follow-up of children beyond pre-school age, with conflicting results. Reports of normal attention and behavior in a small longitudinal study of children conceived with ART [68-70] were followed by an alarmingly high prevalence of self-reported attention deficit hyperactivity disorder (ADHD) in a cross-sectional study of young adults conceived with ART. [71] A couple of large register-based studies on the other hand found children born after ART at lower risk of ADHD compared to all other children.[72,73] Adjustment for maternal factors resulted in a slight reversal of the association in the Swedish study, indicating a weak positive association between ART and drug-treated ADHD.[73] Findings on school performance of children conceived with ART have shown a pattern similar to ADHD. Two register-based studies conducted in Nordic countries looked at children's educational achievements in the national school register at 9th grade.[74,75] Both found children conceived with ART at an overall advantage compared to all other children, [74,75] but findings were reversed to instead show modestly worse school performance when adjustment was made for parental demographic characteristics.

2.3.2 Psychiatric outcomes

Very few studies have evaluated adolescents or young adults, due to the limited availability of follow-up with validated assessments. Since some psychiatric disorders only start to become manifest in adolescence, studies on mental health outcomes like depression and anxiety are extremely scarce. Previous findings have shown concerns that depression and anxiety levels may be higher among children that were conceived with ART.[69,71] These small studies were based on parental and teacher observations. After subsequent follow-up in which the children completed self-questionnaires, no difference was found, suggesting a discrepancy in reporting between parents/teachers and the children themselves, although age may also be a factor since some time had passed between the two assessments.

2.3.3 Childhood asthma

Asthma is one of the most common diseases in childhood. In a nationwide assessment of children in Sweden between 1996 and 2006, 7.4% of 7-8 year old children had received a diagnosis of asthma, and 13% experienced wheezing in the past year.[76] The most common symptom of asthma is wheezing. When exposed to certain triggers, such as contracting an upper respiratory tract infection or inhaling pollen, the airways and lungs become easily inflamed. Childhood asthma is also one of the leading causes of emergency unit visits, hospitalizations, and school days missing during the childhood period. While some grow out of their childhood asthma, the conditions often persist into adulthood.

Children conceived with ART have been found at higher risk of asthma compared to children in the general population (OR 1.20 to 2.65).[77-83] Some have further explored the association stratified by preterm birth and reported a stronger positive association among children born preterm.[79,81] However, the fact that preterm birth is likely an effect of ART was not properly addressed in the studies. A large register-based study conducted on almost half a million individuals in Norway,[11] reported similar findings (RR=1.20, 95% CI 1.0-1.32). They further explored the role of ART itself by controlling for parental infertility and stable parental factors by contrasting the children of couples that had conceived both with and without ART (i.e. sibling comparison).[81] The previously observed association remained, although it was estimated with much poorer precision (RR=1.20, 95% CI 0.85 to 1.71).

2.4 EVALUATION OF ART SAFETY

In the evaluation of the safety of ART, a major challenge is to understand whether differences are due to the effect of the treatment per se or to background factors that increase the likelihood of ART use as well as the outcomes of interest. In case of a causal effect, it is of further interest to understand whether this is due to putative mediators on the causal pathway. (Figure 2.4)[84]

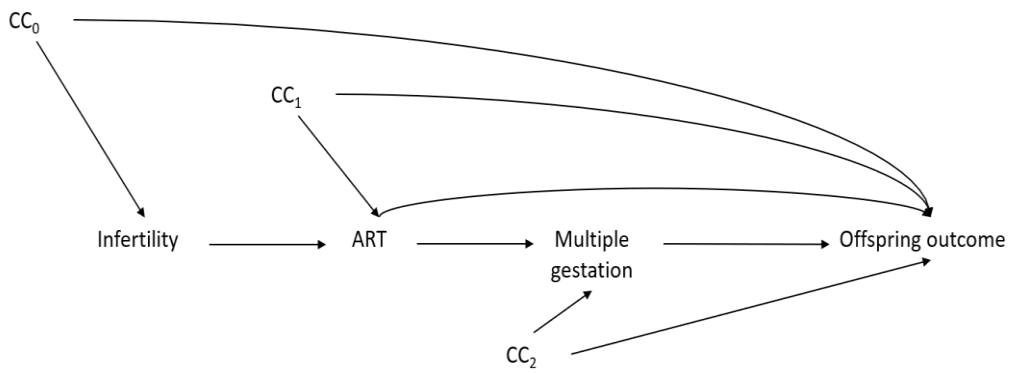


Figure 2.4: Directed Acyclic Diagram (DAG) on ART and offspring outcome

Among all the possible background factors, infertility is the vastly dominant reason couples seek and obtain ART treatment. Evidence suggests that adverse birth outcomes are more common in couples with infertility, also when the children were conceived naturally.[85,86] Also long-term health outcomes appear more common in children of couples with infertility,

including mental disorders [87,88] and childhood asthma.[79] As maternal asthma has been associated with repeated pregnancy loss, a plausible mechanism could be that asthma and infertility share immunological/genetic mechanisms, which would make some children of couples with infertility more susceptible to asthma through genetic covariance.[89-91] Following this reasoning, observed associations between infertility and adverse offspring outcomes are most likely attributed to common causes (CC_0) of the two, ranging from demographic characteristics and medical conditions to genetic factors. Since these common causes give rise to a non-causal association between ART and adverse outcomes (via infertility) it is critical to block this influence in order to isolate the causal influence of ART. To identify and measure all of these various common causes may not be feasible/possible, which is why it is essential to directly account for infertility, as that will block the influence from these (CC_0). Most of the previously mentioned studies [68-71,74,75,78,92] compared children conceived with ART to the general population. While they would also attempt to adjust for some of the factors that influence both ART and offspring outcome (CC_0 and CC_1), few would consider infertility and then typically by adjusting for time to pregnancy.[67,73,77,79,80] To rather account for infertility directly by restriction to children of couples with infertility will not only block all potential influence of risk factors of infertility, but also provide a more meaningful comparison of the at-risk population (having the indication for treatment).

It is plausible that several of the procedures in ART could directly influence offspring long-term outcomes, e.g., embryo and fetal environment on the risk of asthma.[93-95] Since ART can lead to more multiple gestations, and various adverse perinatal outcomes (as previously reviewed) it is also possible that some or all of a potential excess risk from use of ART is a result of these conditions.[96] This classic case of mediation has most commonly been addressed by simple adjustment or stratification on the potential mediators. For instance, many studies on early childhood development were adjusted for preterm birth and low birth weight, and performed in singletons and multiples separately. However, these approaches are problematic as they can introduce bias by opening a non-causal pathway over potential common causes (CC_2 in Figure 2.4) of the mediators and the outcome.[97] For instance, when investigating the association between ART and asthma among children with preterm birth, the non-ART group must have other causes for premature birth, some of which could also influence the risk of asthma, like maternal smoking during pregnancy (SDP). Among preterm children, ART will be inversely related to asthma, and the stratified analysis will thus be biased unless maternal smoking in pregnancy is accounted for.

2.5 GAPS IN THE LITERATURE

At present, evidence regarding the long-term health of children conceived with ART is still lacking, especially such that also evaluate the safety of the procedures. Aside from causal inference, limited follow-up, insufficient sample sizes, and inadequate outcome ascertainment are the major challenges in earlier studies. In the research of neurodevelopmental disorders, the majority of previous studies followed children to a maximum of 5 years. As children grow older, the pattern and expression of neurodevelopmental impairments may change.

Early impairment may not necessarily influence later function, and more complex cognitive, behavioral, and psychiatric manifestations, may not appear until in adolescence or early adulthood. To our knowledge, no previous studies have examined the specific risk of obsessive-compulsive disorder (OCD) or extended the investigation to other indices of mental health, such as suicidal behavior and antidepressant use. We thus identified a need for longer-term evaluation of more consequential endpoints, and careful consideration of causal pathways. Register studies offer unique possibilities for large-scale, nationwide assessment of several indicators of neurodevelopment and psychiatric health across childhood, adolescence, and even into early adulthood.

With the rapid growth of both ICSI and frozen embryo transfer, comes the need to also understand the safety of these specific procedures. Previous studies of ICSI have shown conflicting results. For example, an early evaluation in 5-8 year-olds reported that those conceived with ICSI had a 7-point lower mean IQ than spontaneously conceived children,[98] whereas two other studies found no such differences.[72,99] Comparing ICSI to standard IVF, an early study reported no differences in children's IQ,[99] while another found on average 3 points lower IQ in the ICSI group.[98] Finally, a large Swedish register-based study has noted an elevated risk of mental retardation among children conceived with ICSI.[67] Since ICSI was initially only used for male factor infertility, it has not been possible to disentangle the consequences of the intervention from that of its underlying indication, or from the fact that women undergoing ICSI were generally younger and healthier than the standard IVF group. Concerns for adverse consequences of frozen embryo transfer have been mitigated by an increasing body of studies noting a lower risk of preterm birth, low birth weight, and small for gestational age compared to fresh embryo transfers.[100] However, it has also been linked to an elevated risk of high birth weight and large for gestational age babies.[101,102] Children born after frozen embryo transfer could have a different risk profile also with respect to long-term outcomes. Due to its later implementation, long-term follow-up of children born from frozen embryo transfer is even more scarce than for ICSI. In an early report, no difference was found with respect to the risk of mental retardation compared to fresh embryo transfer.[67] Overall, the evidence regarding the long-term impact of specific procedures such as ICSI and frozen embryo transfer is scarce and suffers the previously mentioned shortcomings (small sample sizes, inadequate length of follow-up). Once again register-based studies allow adequate sample size and length of follow-up to improve our understanding of the safety of these practices.

Clinical prediction of maternal complications during pregnancy may allow more efficient intervention from early detection (such as aspirin for preeclampsia) and monitoring and preparedness of pregnancy progress (such as early delivery, cesarean section, and potential blood loss) for women undergoing ART treatment. At present, the application of machine learning algorithms to ART treatment has been restricted to predicting start day or ovarian stimulation,[103] embryo selection for implantation,[104] early pregnancy loss,[105] clinical pregnancy,[106] and live-birth delivery.[107] In the general population, existing studies have attempted to predict complications in pregnancy using information from medical records,

images, biological markers, ultrasound assessments and fetal heart rate, and mostly relying on logistic regression modeling. [108] Preeclampsia has been the main target for prediction using machine learning methods. Because its multifactorial pathogenesis has not been thoroughly understood, prevention and prediction of preeclampsia are still very difficult.[29] Only one study to date has achieved good performance with an area under the curve (AUC) score of 0.964 and an accuracy of 89%.[109] Another study that applied deep learning methods to predict preeclampsia transcriptome using microarray data obtained an AUC of 0.908.[110] Since women that achieved their pregnancy with ART have been found at elevated risk of these outcomes, it should be of interest to evaluate prediction within this group and specifically consider patient characteristics and treatment-related information.

3 RESEARCH AIMS

The overall aim of this doctoral thesis project was to investigate the long-term health of children conceived with ART with particular focus on the safety of the intervention, and to develop models to predict maternal complications in pregnancy following ART, using large population-based data from Swedish national registers.

The specific research questions for each study were:

- **Study I**: does ART use put children at elevated risk of ADHD and/or lower school performance at 9th grade?

- **Study II**: are children and adolescents conceived with ART at elevated risk of psychiatric disorders, antidepressant use, and suicidal behavior? If so, what is the role of parental characteristics versus the use of ART per se?

- **Study III**: is the reported link between ART and childhood asthma due to the potential influence of confounding or the use of ART specific procedures?

- **Study IV**: can we predict pregnancy and delivery complications in women that conceived with ART? Could the amendment of detailed treatment information and different machine learning algorithms improve prediction performance?

4 MATERIALS AND METHODS

4.1 DATA SOURCE

In Sweden, all residents receive a unique personal identification number at the time of birth or immigration, which allows researchers to track individuals across multiple demographic registers held at Statistics Sweden (SCB, *Statistiska Centralbyrån*) and health care registers held at the National Board of Health and Welfare (SOS, *Socialstyrelsen*).

The Medical Birth Register (MBR, *Medicinska födelseregistret*) was established in 1973 and covers 96-99% of all live births in Sweden.[111,112] The register includes all live births and stillbirths, which means that pregnancies ending in miscarriage or abortion before the 23rd week of pregnancy (29th up to 2008) are not included. Information about pregnancy and delivery are collected from standardized charts used in antenatal care clinics and obstetric units. The information collected by midwives at the first antenatal visit includes maternal weight and height, family situation, tobacco use, reproductive history, and previous/current illnesses such as hypertension and diabetes. The information collected around delivery includes mode of delivery, anthropometric and health status of the newborn, and diagnoses recorded in mother and/or child at the time of discharge using the Swedish version of the International Classification of Diseases (ICD). Since the advent of IVF in 1982, the enrollment interview includes mothers' reports of time to pregnancy, and from 1995 additionally any use of fertility assistance. Up to the end of 2006, the MBR also collected information from IVF clinics in Sweden to record details about the treatment for births achieved with ART.[113]

The National Patient Register (NPR, *Patientregistret*) was established in 1964 and reached nationwide coverage of all hospitalization in Sweden in 1987. In 2001 the register was expanded to also include outpatient visits to specialist care.[114] The register does not include primary care visits. Recorded information includes date of visit or admission and discharge, type of medical specialty, and any diagnoses and procedures made. Diagnoses are recorded using the period-specific Swedish version of the ICD coding system (ICD-7 1964-1968, ICD-8 1969-1986, ICD-9 1987-1996, ICD-10 1997-present), and medical or surgical procedures have been recorded using the Swedish classification of medical procedures *(TAL, Klassifikation av vårdåtgärder, KVÅ).*

The Swedish Prescribed Drug Register (PDR, *Läkemedelsregistret*) was established in July 2005 to record information on all dispensations of prescribed medications at pharmacies in Sweden. [115] Information includes date of prescription and dispensation, dose and amount prescribed, and both the brand name and classification of the medication according to the Anatomical Therapeutic Chemical (ATC) classification system. Over-the-counter purchases of drugs are not included in the register.

The Multi-Generation Register (MGR, *Flergenerationsregistret*) records the biological and adoptive parents for all individuals born from 1932 and who were alive and legally residing in Sweden at any point from 1961.[116] The register allows the creation of family pedigrees across several generations.

The Register of the Total Population (RTB, *Registret över totalbefolkningen*) was established in 1968 and records the date of birth, sex, country of birth, dates of immigration and emigration, and civil status.[117]

The Cause of Death Register (CDR, *Dödsorsaksregistret*) was established in 1961 to record dates, and main and contributory causes of death for all deaths from 1952 onward.

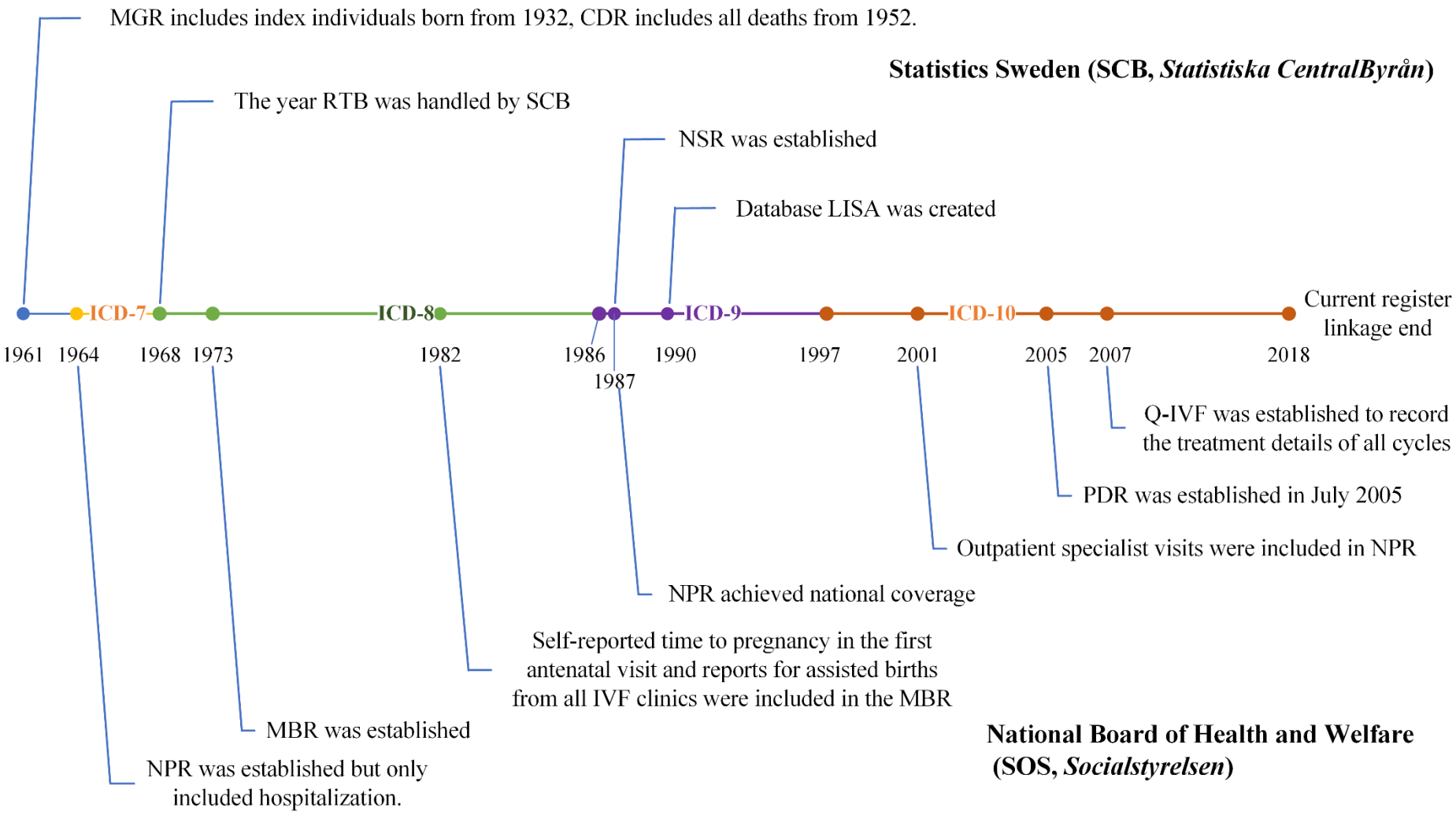
The Longitudinal Integrated Database for Health Insurance and Labor Market Studies (LISA, *Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknads-studier*) was initialized in 1990 and collect detailed data on occupation, the highest level of attained education and individual and family income for all persons aged 16 or older at the end of each year from various registers at SCB, such as the Income and Taxation Register and the Education Register.[118] From 2010, the LISA also includes data on persons aged 15 years.

The National School Register (NSR) was established in 1987 and includes subject grades and scores on national tests in Swedish, mathematics, and English from the 9th grade (end of the compulsory lower secondary school). The information also includes an assessment of eligibility to upper secondary school. The Swedish grading system has changed several times across the covered period.

The National Quality Register for Assisted Reproductive Techniques (Q-IVF, *Nationellt kvalitetsregister för assisterad befruktning*): was started in 2007 when all IVF clinics in Sweden were required to record details about their practice. In contrast to the information previously reported to the MBR, the Q-IVF thus includes details on all performed treatments, irrespective of their outcome with respect to potential pregnancy and birth. The information includes fertilization method (ICSI, IVF, or artificial insemination) and for all ART cycles whether it was a fresh or frozen embryo transfer, whether sperm or oocytes were own or donated, what stimulation protocol the woman underwent, the number of oocytes retrieved, number of embryo culture days, number of embryos transferred, and potential cryopreservation method of embryos.

A timeline of the described registers is provided in Figure 4.1.

Figure 4.1: Summary of registers and ICD systems used in the studies of the thesis


MGR includes index individuals born from 1932, CDR includes all deaths from 1952.
Statistics Sweden (SCB, Statistiska CentralByrån)
The year RTB was handled by SCB
NSR was established
Database LISA was created
ICD-7
ICD-8
ICD-9
ICD-10
Current register linkage end
1961
1964
1968
1973
1982
1986
1987
1990
1997
2001
2005
2007
2018
Q-IVF was established to record the treatment details of all cycles
PDR was established in July 2005
Outpatient specialist visits were included in NPR
NPR achieved national coverage
Self-reported time to pregnancy in the first antenatal visit and reports for assisted births from all IVF clinics were included in the MBR
MBR was established
NPR was established but only included hospitalization.
National Board of Health and Welfare (SOS, Socialstyrelsen)

4.2 STUDY DESIGN

The long-term follow-up of children conceived with ART was performed in birth cohorts identified in the MBR. Outcome follow-up was available through the end of 2018, and different birth periods were considered to allow the capture of the outcomes with respect to (a) register coverage and (b) age of onset. Children born following use of ART were compared to all other "spontaneously conceived" children in the population first, then restricted to children born to couples with infertility, and finally to each other (contrasting specific ART procedures). In Study III, within-family comparisons (full siblings and first-born maternal cousins) were also explored.

The prediction of maternal complications during pregnancy and after delivery was performed in a cohort of nulliparous women undergoing ART, identified in the Q-IVF. The study was restricted to women that achieved birth following their first, second, or third attempt (defined as consecutive embryo transfers) respectively.

Table 4.2: An overview of the individual studies in the thesis

Study	Study design	Study period	Study population	Exposure/ predictor	Outcome
I	Birth cohort	1996-2012	1,655,873 live births	ART, IVF, ICSI, frozen or fresh embryo transfer	ADHD
I	Birth cohort	1986-2001	1,555,573 live births	ART, IVF, ICSI, frozen or fresh embryo transfer	School performance
II	Birth cohort	1994-2006	1,221,812 live births	ART, IVF, ICSI, frozen or fresh embryo transfer	Major depression, anxiety, antidepressant use, suicidal behavior
III	Birth cohort	2005-2013	950,953 live births	ART, IVF, ICSI, frozen or fresh	Childhood asthma

				embryo transfer	
IV	Prediction cohort	2008-2016	17,365 nulliparous women giving birth with ART	Demographic, Medical history, medication, ART treatment information	Preeclampsia, Placental complications (accreta, previa, abruption), PPH

4.3 MEASURES, ASCERTAINMENT, AND DATA QUALITY IN STUDY I - III

4.3.1 Infertility

Information on infertility was critical to identify the primary population "at-risk" of ART use. Children born to couples with infertility were identified through any of the following sources: the mother's self-report at the first antenatal visit; recorded diagnosis of parental infertility in the MBR or NPR; and clinical report of ART use in the MBR from 1982 to 2006, and in the Q-IVF from 2007.

Maternal self-report:

In 1982, a question about involuntary childlessness was added to the standardized interview at the first antenatal visit, with the answer recorded in number of years. Since 1995 a positive response is followed by a question if any measures were taken to overcome this, with the options: no measure, IVF, ovulation stimulation, surgery, ICSI (added in 1999), or other (Figure 4.3.1).

Figure 4.3.1: Questions regarding infertility and fertility assistance in the standardized interview at enrolment to antenatal care

The pink frame indicates that the information is reported to the MBR

Ofrivillig barnlöshet, antal år: involuntary childlessness, number of years

Measure (Åtgärd) in sequence are: nothing (ingen), IVF (IVF), ovulation stimulation (ovul.stim), surgery (kirurgi), ICSI (ICSI), and other (annat:).

Diagnosis of infertility:

Diagnosis of infertility was considered when a clinical record of such was identified in the NPR anytime up to the birth for primiparous parents and anytime from prior to current birth for multiparous parents, or if a diagnosis had been recorded in the MBR. Table 4.3.1 lists the infertility diagnoses that were considered, and also shows how pre-existing conditions related to infertility were identified:

Table 4.3.1: Codes used to identify diagnosis of infertility, infertility evaluation, and infertility-related pre-existing conditions

		ICD-8	ICD-9	ICD-10
Infertility	Female factor	628	628	N97
	Male factor	606	606	N46
	NOS*		V23A	Z35.0, Z31.0, Z31.1
Pre-existing conditions	PCOS**	256.9	256E	E28.2
	Endometriosis	625.3	617	N80

*NOS: not otherwise specified. **PCOS: Polycystic ovary syndrome.

Since the use of ART for other reasons than infertility was quite limited in the birth periods under study (becoming available to women in same-sex couples from 2005), ART use was assumed to indicate infertility in these studies. Still, the primary objective for identifying infertility was to obtain a relevant, at-risk but not exposed, comparison group to the children conceived with ART. To this end we considered also those that had undergone infertility evaluation (Z35.0) and surgery/artificial insemination (Z31.0, Z31.1), along with those reporting more than 1 year of involuntary childlessness in the MBR (capturing also couples who had not sought help to evaluate and/or overcome their infertility).

4.3.2 ART

Similar to infertility, use of ARTs was identified by combining information from three principally different sources. The primary source was IVF-clinic reporting (MBR or Q-IVF) matched to the birth, followed by diagnostic or procedure codes for ART use in the NPR within 1 month of the estimated date of conception, and self-reported use of IVF or ICSI in the MBR. Table 4.3.2 shows the ART diagnostic and procedure codes in Study I-III:

Table 4.3.2: Codes used to identify a diagnosis or procedure involving IVF and ICSI

		ICD-10	Procedure code (*KVÅ)*
ART	IVF	O26.8A, Z31.2, Z31.3, Z31.2A, Z31.2C	LCA30, DL002, DL003
	ICSI	Z31.2B	DL001, DL007

All three sources allowed identification of potential use of ICSI (as opposed to standard IVF), whereas only the IVF-clinic reporting included information on whether the embryo transfer was fresh or frozen.

4.3.3 Outcomes

ADHD:

In the ADHD cohort of Study I, children born between 1996 and 2012 were followed until whichever occurred first of ADHD, emigration, death, or the end of 2018. The cohort was selected to allow diagnostic coverage from at least age 5. In Sweden, the diagnosis, treatment, and monitoring of ADHD are all handled by specialist physicians.[119] Identification of ADHD was considered using an established algorithm [120,121] which has also been previously validated.[121] The algorithm was based on records of diagnosis (ICD-10: F90) in the NPR and/or dispensation of ADHD medication in the PDR (ATC: N06BA01, N06BA02, N06BA04, N06BA09, N06BA12, C02AC02). The first record, either from diagnosis or dispensation, was considered the incident date of diagnosis.

School performance:

The school performance cohort of Study I included all children born between 1986 and 2001, to ensure they would have completed the 9th grade by 2018. Measures of performance at the completion of compulsory (lower secondary) school included: 1) the final average grade (0–20) which was the mean grade across a total of 16 subjects, with each subject grade ranging from 0 to 20; 2) eligibility for upper secondary school which was a binary measure for passing in at least 12 subjects, including Swedish, English, and mathematics.

Psychiatric disorders:

In Study II, children born between 1994 and 2006 were followed until whichever came first of the event of interest, emigration, death, or the end of 2018. The interval was selected to allow full coverage of outpatient specialist visits from a minimum age of 7 (2001) and to allow everyone to be followed until at least age 12. Up to and including the age of 17,

diagnosis and treatment of psychiatric disorders are typically handled by specialists in child and adolescent psychiatry. Once the individual has turned 18, they may be diagnosed and treated by general practitioners in primary care, or by specialists in psychiatry. The former would typically refer more complicated or severe cases to the latter. In addition to a broad capture of mood (ICD-10: F30-39) and anxiety disorders (ICD-10: F40-42, F93) we specifically investigated major depression (ICD-10: F32, F33) and stratified anxiety disorders into obsessive-compulsive disorders (OCD) (ICD-10: F42) and non-OCDs (ICD-10: F40, F41, and F93). We also used dispensation of antidepressant medication (ATC: N06A) among those with a depression diagnosis, as an indicator of more severe depression. Any dispensation of antidepressant medication was also used alone, as a broader indicator of psychiatric morbidity. Use of antidepressants (ATC: N06A) was identified through the PDR. The medications were further grouped into selective serotonin reuptake inhibitors (SSRI) (ATC: N06AB) and all other (non-SSRI antidepressants). Lastly, we also considered suicide, including completed suicides from death certificates in the CDR (ICD-10: X60-X84) and suicidal attempts of determined intent (ICD-10: X60-X84) from patient records in the NPR.

Childhood asthma:

Study III included all children born between 2005 and 2013 and followed them up until incident asthma, emigration, death, or the end of 2018, whichever came first. Clinically, asthma is defined as "3 or more obstructive periods before the age of 2 and/or 1 or more obstructive periods after the age of 2, and/or 1 or more obstructive periods independent of age when the child has 1 or more of the following: eczema, allergy or no improvement between periods of respiratory tract infections." [122] In this study, asthma was identified based on a validated algorithm, combining records of diagnosis and dispensed drugs:[122,123]

At least one clinical diagnosis (ICD-10: J45-46) and either of the following dispensation algorithms (ATC: R03):

(a) ≥2 dispensations of inhaled corticosteroids, leukotriene receptor antagonists, or fixed combinations of β2 agonists and corticosteroids, with gaps of at least 2 weeks, or

(b) ≥3 dispensations of the above-mentioned drugs or short-acting β2 agonists within 1 year.

Incident asthma was defined from the first available record contributing toward the criteria of the algorithm. To tease out more persistent disease, current asthma was defined as meeting the criteria for incident asthma followed by at least one diagnostic or dispensation record at a specific age, namely, within the year the child was 3, 4, 5, or 6 years old respectively.

4.3.4 Other variables

In the evaluation of long-term safety of ART, we considered a wide range of factors that may directly (CC_1 in Figure 2.4) or indirectly (CC_0 in Figure 2.4) impact both the parental use of ART and the risk of the outcomes of interest. These were identified and generally considered in broader groupings. First, parental demographics included age at child's birth, country of origin, and the highest level of attained education. Mother's country of origin was obtained

from the MBR whereas the father's country of origin was obtained from the RTB, and both were grouped to Sweden, other Nordic countries, other European countries, and non-European countries. Parental education was obtained from the LISA and categorized as lower secondary (9 years or less), upper secondary (9-12 years), and post-secondary (>12 years). Second, maternal characteristics at the start of pregnancy included parity, civil status, smoking in early pregnancy, overweight or obesity, and region of residence. Maternal smoking in early pregnancy was self-reported at the first antenatal visit and was originally recorded in 3 levels: no smoking, 1-10 cigarettes per day, and more than 10 cigarettes per day. Overweight or obesity was defined as body mass index (BMI) greater than 25, calculated from the mother's reported height and weight at the first antenatal visit. Since clinical practice, e.g., the use of ART treatment and diagnosis of the outcomes under study, could vary across the country, the mother's county of residence in the year of the birth was categorized into five regions (North, South, Mid, East, and West). Third, maternal health history included but was not limited to the infertility-related conditions PCOS and endometriosis. Forth, parental medical history specific to each outcome of interest. For example, in Study II we considered parental psychiatric history as a summary indicator for couples who had any diagnostic record of a mood disorder or a non-affective psychosis, and in Study III we considered parental history of asthma. Finally, birth-specific characteristics such as multiple gestation, preterm birth, and birthweight were also available, and considered to be possible mediators in the causal pathway (more below).

4.4 EPIDEMIOLOGICAL METHODS IN STUDY I - III

4.4.1 Cohort study design

The goal of a cohort study is to measure and often to compare the incidence of one or more diseases. The word "cohort" designates a group of people who share a common experience or condition, such as exposure. In a typical cohort study, individuals with and without a particular exposure are followed prospectively with respect to the potential occurrence of the outcome of interest. Depending on the data source used, some studies rely on primary data collection where the researcher defines what to measure. This poses logistical and financial challenges, especially when it takes long for the outcome to develop. However, this is not an issue in studies using pre-existing data sources where the information was routinely collected and primarily intended for clinical documentation or administrative purposes but then can be secondarily used for research, such as nationwide registers in the Nordic countries. Also, in studying treatment effects, it can be challenging to pinpoint the start of the induction time from exposure to outcome. This was not a problem in the studies of the thesis, as the intervention of ART is confined to a limited and well-defined time point (around conception).

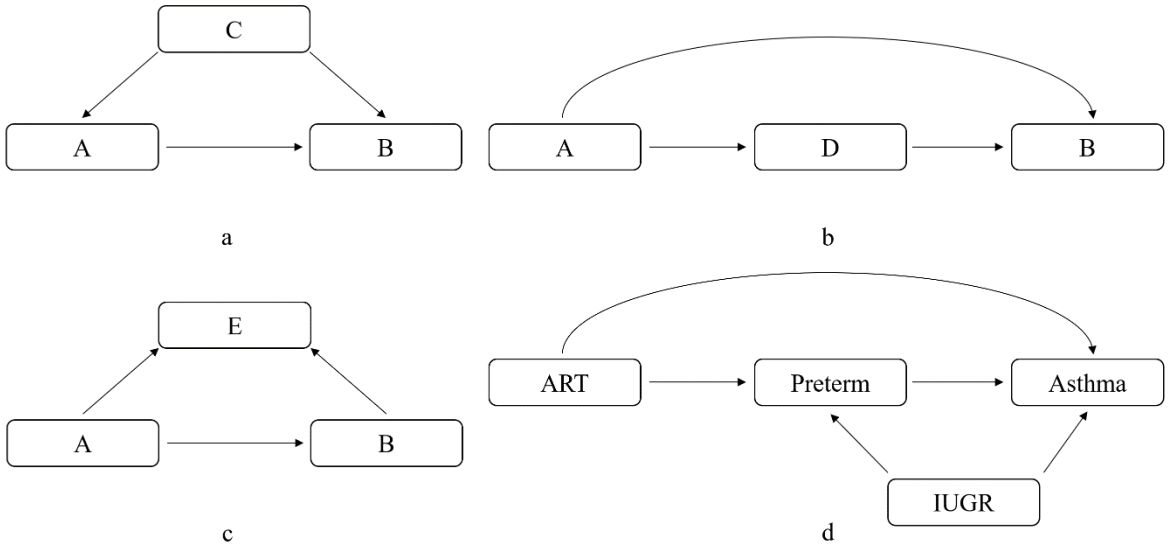
4.4.2 Confounder, mediator, and collider in causal diagram

Confounding is a central consideration of causal inference. In almost all non-randomized studies, the exposed and unexposed will be systematically different not only in their exposure and some of the factors that influenced the likelihood of exposure may also influence the risk

of the outcome. The presence of such common causes of exposure and outcome creates a spurious association between the two which is called confounding, A confounder is any factor that could block such confounding influence. In Figure 4.4.2a we aim to study the effect of component A on B, and component C influences the distribution of both A and B, i.e., a common cause. The observed association between A and B can be due to a causal effect of A on B, but also the open path over C. If we hold C constant it has no influence, hence the common cause C is a confounder. While all common causes are confounders the converse is not the case. In figure 2.4 for example, infertility is a confounder since when we hold it constant, there is no confounding influence from CC_0, even though infertility is not itself a common cause of ART and the outcome. If confounding is present, we say the association of an exposure with an outcome is confounded, or biased. Whenever there is a systematic error in what we wish to measure – the causal effect – this can be described as bias. Bias produced by confounding can be in any direction, i.e., it can lead to underestimation or overestimation of the causal effect. The random assignment of treatment in randomized controlled trials ensures that there are no systematic differences between exposed and unexposed, such that there can be no confounding. In observational studies, we need to identify all systematic differences that could also influence the outcome using subject-matter knowledge. Directed acyclic graphs (DAGs) can then help identify how the influence from common causes can be blocked. This is achieved through designs and methods that ensure the comparison is made among individuals that are similar with respect to the confounders, such as matching, stratification, and standardization. In our studies, we used a form of stratification by adjusting for measured confounders in regression and restricted the comparison to groups with similar characteristics.

A mediator is a factor that is on the casual pathway from the exposure to the outcome (D in Figure 4.4.2b). The indirect effect of exposure on the outcome is blocked when the mediator is held constant (stratified on). This could be problematic if the mediator turns out to be a collider, i.e., a common effect, on a path between exposure and outcome, since holding a collider constant opens the path over it (that was otherwise blocked) (E in Figure 4.4.2c). An example is also shown in Figure 4.4.2d. The association between ART and asthma could be mediated by preterm birth. Intrauterine Growth Restriction (IUGR) is a cause of the mediator preterm birth and the outcome asthma. To assess the direct effect of ART on asthma (not mediated by preterm birth) we would like to adjust for preterm birth. However, unless we also adjust for IUGR a new pathway will be opened from ART through preterm birth and IUGR to childhood asthma. This is because, particularly in the preterm birth stratum, the non-ART group are more likely to be preterm due to IUGR than the ART group, and thus in turn also more likely to develop asthma. The resulting non-causal relation between ART and asthma is sometimes called collider-stratification bias, and is a form of selection bias.

Figure 4.4.2: Illustration of confounder, mediator, and collider



4.4.3 Family-based study design

It is rarely feasible to be able to measure all confounders required to block the influence of common causes, and family designs can then help as one can achieve control for all factors shared in families without having to measure them. This is especially useful when suspecting genetic confounding since rarely can we measure genetic factors. The principal idea of family-based designs is to compare exposure discordant relatives such as twins, siblings, and cousins. Due to the varying degree that different types of relatives are similar, the choice will influence the type and extent of confounding control. Besides the automatic adjustment for all shared confounders, it should be noted that all shared mediators on the casual pathway will also be held constant. This could be an issue if the total effect of the exposure is the aim of the analysis. It is also essential to know that bias due to measurement error and non-shared confounders may be inflated when the comparison is restricted to exposure-discordant relatives.[124] Another important assumption in sibling comparison designs is that there should be no carry-over effect in siblings.[125]

4.5 STATISTICAL METHODS IN STUDY I - III

When quantifying the difference in outcome between different groups (exposed and unexposed) we need to take into account the influence of the random error that arises from sampling variability, by using statistical methods that inform us about the precision of our estimation. When further seeking to quantify whether an association reflects a causal effect, we need to take into account the influence of systematic errors, such as confounding, by using statistical methods that allow us to hold confounders constant. In Study I-III, different statistical methods were considered because of the type of study design and outcome.

4.5.1 Generalized linear models

When the outcome is a continuous measure, like the average grade in Study I, the association to ART can be evaluated using multivariable linear regression, as shown in equation 4.5.1a.

$$Mean\ score\ (\mu) = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon \quad \textbf{Equation 4.5.1a}$$

The basic principle is to let the mean outcome μ (average grade) be predicted by a set of covariates X_1-X_n (ART and potential confounders). The intercept β_0 is the mean value of the average grade when all covariates are at the reference level (Vector X is set to 0). Each coefficient (β_1~β_n) is interpreted as the change in the mean outcome per unit change of the corresponding covariate X, while holding all other covariates constant. The error term ε represents all the variability in the mean outcome that cannot be explained by the vector of covariates (X_1-X_n).

Linear regression can be generalized to accommodate other types of outcomes, whereby a link function transforms the outcome Y such that it can be related to the linear predictor (Link (Y)= $\beta_0 + \beta_1 X_1 + \varepsilon$). All generalized linear models (GLM) operate under the following assumptions: 1) A linear relationship between X and Y; 2) No multi-collinearity between the covariates (X_1-X_n) in the model; 3) The residuals (the difference between observed and predicted outcome) are normally distributed; 4) The residuals are independent, and finally 5) the variance of the residuals is constant across all values of X (homoscedasticity).

When the outcome is binary, like the eligibility for upper secondary school in Study I, the mean outcome corresponds to a probability. Since probabilities are constrained to 0-1, the generalized linear form can be maintained with the logit (log odds) link function, which is mathematically more convenient to model, shown as equation 4.5.1b:

$$logit\ p(x) = ln(\frac{P(pass = 1)}{1 - P(pass = 1)}) = \beta 0 + \beta 1 X \qquad \textbf{Equation 4.5.1b}$$

With some transformation, the probability of obtaining eligibility for upper secondary school can be estimated, shown as equation 4.5.1c:

$$\frac{P(pass = 1)}{1 - P(pass = 1)} = e^{(\beta 0 + \beta 1 X)} \qquad \textbf{Equation 4.5.1c}$$

$$P(pass=1) = \frac{e^x}{1+e^x} = \frac{e^{(\beta 0+\beta 1 x)}}{1+e^{(\beta 0+\beta 1 x)}}$$

Due to the logit link, the coefficient $\beta 1$ represents the difference in log-odds (odds being probability of passing versus the probability of not passing) and once exponentiated, the coefficient for a binary exposure variable (X=1 or 0) can be interpreted as an odds ratio (exp ($\beta 1$)) (equation 4.5.1d):

$$\beta 1 = ln\left(\frac{p(1)}{1-p(1)}\right) - ln\left(\frac{p(0)}{1-p(0)}\right) \qquad \textbf{Equation 4.5.1d}$$

$$= ln\left(\frac{\left(\frac{p(1)}{1-p(1)}\right)}{\left(\frac{p(0)}{1-p(0)}\right)}\right) = ln\,(\frac{odds\ of\ passing\ among\ X=1)}{dds\ of\ passing\ among\ X=0)})$$

$$=ln\ (Odds\ Ratio)$$

$$OR = exp(\beta 1)$$

Logistic regression can be very handy because the log-odds transformation of the mean outcome allows a more stable prediction from any value (from negative to positive infinity), such that it can be predicted by a linear combination of the covariates (with assumptions of linear regression). The price we pay is that the effect measure reflects the less intuitive odds ratio (rather than a risk ratio).

4.5.2 Survival models

When the outcome is the occurrence of an event, such as with the different conditions in Study I-III, we typically cannot observe the outcome in all individuals due to loss of follow-up (in register-based studies primarily because of emigration), death, or because the study follow-up ends. Survival models are designed to account for censoring, i.e., when individuals no longer are informative with respect to risk of the outcome. The most common method to model time-to-event data is to model the instantaneous hazard (event rate) by using Cox proportional hazard regression. A simplified version of Cox regression is written in equation 4.5.2. The Cox model conveniently avoids modeling of the baseline hazard $\lambda_0(t)$, such that it is semi-parametric, but instead rests on the assumption that the hazards are proportional over time.

$$\lambda(t|X) = \lambda_0(t)\exp\,(\beta X) \qquad \textbf{Equation 4.5.2}$$

From this follows that Cox proportional hazard models are invalid when the hazard varies across strata of a covariate. However, there are several solutions to such potential violations. Stratified Cox regression allows the baseline hazards to vary in strata of one covariate. The regression can also be performed in separate time bands or include the use of time-varying covariates (allowing interaction with time). Alternatively, flexible parametric survival regression models can be used to model the baseline hazard flexibly using restricted cubic splines. In Study III we found that the relationship between ART and asthma varied with age,

hence we explored the association using both flexible parametric models and stratified Cox regressions with time-bands.

4.5.3 Multiple imputation

In Study I-III, we encountered missing information in some of the covariates of interest, with the most affected being maternal BMI, smoking in early pregnancy, civil status, and region of residence. Missing data can be characterized according to three principle mechanisms for the missing; missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR).[126] In contrast to MCAR in which the missing mechanism has no relation to the inherent values or other characteristics, MAR means that the missing is systematic, but this can be accounted for by other observed variables. Finally, MNAR means data is missing for reasons related to the values themselves. Imputation of missing information can be done for both MCAR and MAR, however, the assumption of independence from the values themselves cannot be tested. The type of imputation can vary from simple replacement of the nearest value or the sample mean, to lending information from other characteristics through multiple imputations. In Study I-III, we imputed multiple datasets using the fully conditional specification (chained equation) method,[127,128] which is a development from Bayesian simulation. Using Rubin's rules,[129] the final estimates were obtained from the results of each imputed dataset. When selecting variables to predict the values for imputation it is important to consider also the outcome, since its omittance could lead to attenuation of the associations between exposure and outcome.[126]

4.6 OUTCOMES AND PREDICTORS IN STUDY IV

Contrary to the studies of ART safety, in which the ultimate goal is to understand the extent observed associations are due to causal influence or non-causal explanations (confounding), Study IV aimed to see the extent of background and treatment characteristics that could predict the risk of maternal complications in pregnancies achieved with ART, and identify the most influential factors.

The outcomes of interest in the prediction models were pre-eclampsia (ICD-10: O14, O15), placental complications including placenta accreta (ICD-10: O43.2), placenta previa (ICD-10: O44), and placental abruption (ICD-10: O45), and PPH (ICD-10: O72) identified from diagnosis made in the MBR or the NPR. In addition to the diagnoses recorded in the MBR (from the epicrisis following delivery), we considered any admission/visit in the NPR during pregnancy for preeclampsia and placental complications, or discharges within 7 days of delivery for PPH.

To investigate prediction in the specific population of women who conceive with ART, the probability of each of the complications of interest was predicted at two clinically distinct time points; first before the initiation of the ART attempt (leading to embryo transfer), and second once pregnancy had been achieved. Since the number of attempts required to achieve

birth varies between women, the predictions were made separately for women giving birth following the first, second, and third attempt respectively.

To that aim, we considered a wide range of potential predictors focusing on information that would be available at the start of treatment and at enrolment in antenatal care respectively. The characteristics considered for the "pre-treatment" model included:1) demographic factors (year, age, partner's age, education, country of origin, BMI, civil status, and region of residence); 2) infertility-related medical history (female or male infertility diagnosis, ovarian dysfunction, endometriosis, etc.); 3) other medical history (hypertensive disease, diabetes mellitus, heart disease, rheumatological disorders, thyroid disease, and coagulopathy); 4) use of contraceptives and/or medication to induce ovulation in the 3-year period prior to ART; and 5) planned ART protocol details (stimulation protocol, own or donated egg and sperm, and fresh or frozen embryo transfer in the 2nd and 3rd attempts). For the "post-treatment" model we additionally considered details about the implemented ART protocol (total gonadotropin dose received, number of eggs retrieved, and for the transferred embryo(s) additionally the fertilization method, stage, number, and potential cryopreservation method), along with maternal self-report of tobacco smoking and time to pregnancy from the first antenatal visit. Table 4.6 shows the identification of predictors used in Study IV.

Table 4.6: Potential predictors for the pre- and the post-treatment models

Predictors	Timing	ICD-10/ATC codes or assigned levels	Pre-treatment model	Post-treatment model
Demographics				
Age at treatment	At each ART cycle		✓	✓
Partner age at treatment			✓	✓
Highest level of education			✓	✓
BMI			✓	✓
Civil status			✓	✓
Region of residence	Latest prior record		✓	✓
Country of origin	At each ART cycle		✓	✓
Cycle year			✓	✓
Medical history (Infertility-related conditions)				
Female infertility	Within a maximum of 10 years prior to ART	N97	✓	✓
Male infertility		N46	✓	✓
Ovarian dysfunction		E28	✓	✓
Endometriosis		N80	✓	✓

Medical history (Other medical conditions)				
Hypertensive disease	Within a maximum of 10 years prior to ART	I10-I16	✓	✓
Diabetes mellitus		E08-E13	✓	✓
Heart disease		I20-I25, I30-I52	✓	✓
Rheumatological conditions		M05, M06, K51, K52, M32	✓	✓
Thyroid disease		E00-E06	✓	✓
Coagulopathy		D68	✓	✓
Medical history (Prescribed medication use)				
Ovulation induction	From 3 years to 3 months prior to the first ART cycle	G03G, L02BG	✓	✓
Contraceptives		G02B, G03A	✓	✓
ART treatment information				
Type of embryo-transfer	In each ART cycle	fresh/frozen	✓*	✓
Stimulation protocol			✓	✓
Total gonadotropin dose			×	✓
Egg origin		own/donation	✓	✓
Number of eggs			×	✓
Sperm origin		own/donation	✓	✓
Stage of the transferred embryo(s)		cleavage / blastocyst	×	✓
Type of fertilization		IVF/ICSI/combination	×	✓
Number of transferred embryos			×	✓
Type of cryo-preservation		slow freezing/	✓*	✓
Number of fetuses			×	✓
Early pregnancy information from the first antenatal visit				
Maternal smoking	Pregnancy record matches the QIVF record		×	×**
Self-reported time to pregnancy			×	✓

*: only in cycle 2 and cycle 3; ** was dropped in the main analyses.

4.7 MACHINE LEARNING METHODS IN STUDY IV

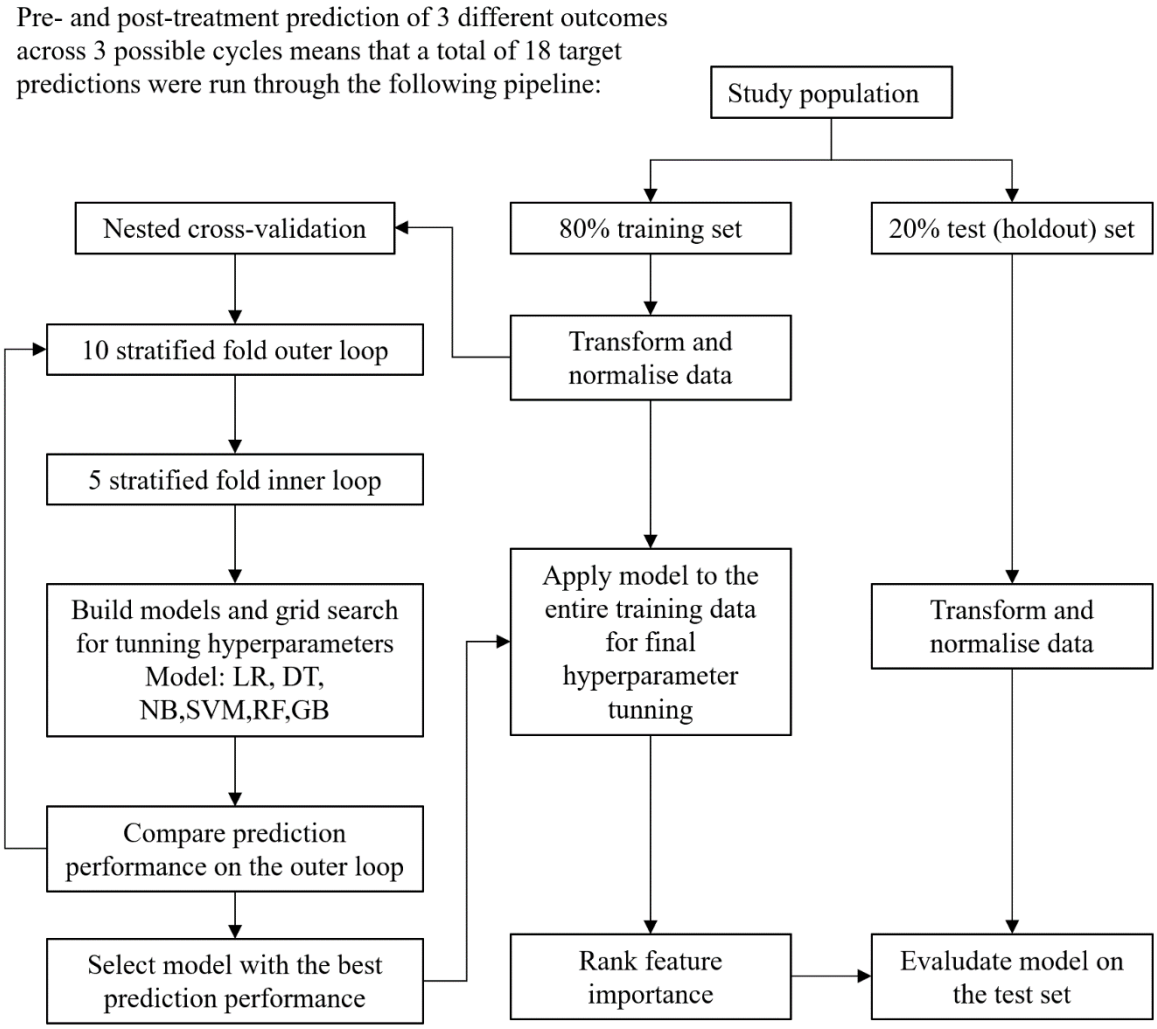
4.7.1 Training set, validation set, test set, and cross-validation set

Pre- and post-treatment prediction of the three outcomes of interest and for the first, second or third ART attempt separately yielded a total of 18 target predictions that each underwent a pipeline of model development, using a machine learning approach.

Machine learning is often used to make predictions based on data, and the data is also the source for the machine to learn and generate model algorithms. Three datasets for training, validation, and testing respectively are needed as input data when creating the model. The first is used to train a model by fitting the parameters of predictor or classifier, using for example logistic or linear regression. The validation set is used to tune the hyperparameters of a classifier. Unlike parameters, hyperparameter refers to a special parameter (often artificially created) whose value is used to control the learning process. An example of a model algorithm and its hyperparameter is the "least absolute shrinkage and selection operator" (LASSO) which introduces a regularization/penalty hyperparameter to ordinary least square regression. Finally, the test set is the data set on which the model will be applied after training and validation. If possible, the test set should be able to represent the training and validation sets while being independent of the two. The test set is sometimes also called "holdout", as it provides an estimate of the learning performance of the machine learning model and should only be used once. The holdout set is not designed for choosing or improving algorithms. In the learning process, the training and validation sets can be decided with pre-defined portions, or they can be used with different combinations of the data to test and train a model on different iterations. The latter is referred to as cross-validation, among which the most common type is K-fold. This is a data partitioning strategy that requires the data to be split randomly into k groups. When a specific k is chosen, one group is treated as a validation set while the other k-1 is used as training data. In situations when the data has an imbalanced classification problem, typically due to the rare occurrence of a clinical outcome, stratified K-fold cross-validation is often used to ensure that each fold of the dataset has the same proportion of observations with the outcome of interest.

In Study IV, we further considered nested cross-validation. The nested cross-validation includes two-fold loops (outer and inner loop) so it was also called double cross-validation. In the inner loop, the k1-fold cross-validation procedure used for choosing the best model by hyperparameter optimization is nested inside the k2-fold cross-validation procedure for model selection in the outer loop. The learning process of machine learning in Study IV is shown in Figure 4.7.1.

Figure 4.7.1: The learning process of predictive models in Study IV, using a machine learning approach



4.7.2 Machine learning algorithms

In machine learning, algorithms refer to all techniques spanning supervised, semi-supervised, unsupervised, and reinforcement methods that analyze input data to predict output values with a closely or loosely controlled function. Since the goal of Study IV was to construct models to determine which category of the outcome individuals belong to, only supervised classification algorithms are discussed in this section. In classification tasks where a conclusion for categorization is drawn from a predicted probability, the top commonly used algorithms are logistic regression (LR), decision tree (DT), naïve Bayes classification (NBC), support vector machine (SVM), random forest (RF), and gradient boosted decision trees (GB).[130-134] LR has been the traditional function to predict binary outcomes. In Study IV, a regularization term was added to the LR. DT is a flowchart-based function in which each branch node represents the conditions towards the leaf node at the branch end, and each leaf node represents a class category, decided after model construction. The hyperparameter for the DT in Study IV was the maximum depth of the tree. Based on Bayes' theorem, NBC is a member of the posterior conditional probability family with predictors strongly independent

from each other. While it was included in the set of algorithms used in Study IV, NBC may not be appropriate given that many predictors were correlated (a suspicion that eventually was supported by consistently poorer prediction performance compared to the other algorithms used). SVM constructs data points over the N-dimensional space (N being the number of input variables) and aims to find a hyperplane that distinctly classifies the data points, using computation from predictors. When using cross-validation, SVM can output the probability of the class label. There were no hyperparameters to tune in the NBC and SVM algorithms in Study IV. RF and GB are different classification methods to combine many decision trees by building each individual tree iteratively and creating an uncorrelated forest of trees whose prediction by ensemble, in principle, is more accurate than that of any individual tree. As such, these ensemble methods use the "wisdom of the crowd". Besides the maximum depth of each tree, the hyperparameter for RT and GB in Study IV also included the number of trees, and in GB additionally also the learning rate was between 0.1 to 1. For hyperparameter optimization, grid search is the most commonly used method for all algorithms used in Study IV.

4.7.3 Measures and interpretation for prediction performance

Measures for prediction performance include accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and receiver operating characteristics (ROC) curves. [135] The ROC curve is a visualization of the prediction of a binary outcome that indicates the tradeoffs between sensitivity and specificity. In Study IV, we measured prediction performance using a score generated for the area under the curve (AUC) of the ROC curves, since outcomes were rare and the predictor is of probabilistic type. Also referred to as C-statistic, the AUC score is one of the most popular measures of goodness of fit for predictions. Ranging between 0-1, with 1 being perfect classification, an AUC score value of 0.5 corresponds to classification made by chance. It is possible that the AUC score is even smaller than 0.5 if the constructed model is no better (or even worse) than a random guess.

4.7.4 Feature importance

Feature importance refers to a technique that, for a given model and a list of input variables (variables are also often called features), calculates a score for each feature to represent how important each feature is. For some machine learning algorithms, the feature importance can be extracted easily. For example, in logistic regression, the magnitude of the regression coefficients can be directly regarded as an indicator of feature importance. However, comparison of feature importance across different algorithms is often non-intuitive, and the feature importance from some algorithms may be more prone to bias.[136] Originating from game theory, the Shapley value has become a universal solution for various methods to assess feature importance.[137] Essentially, the Shapley value is the average expected marginal contribution of one feature after all possible combinations have been considered. SHapley Additive exPlanations (SHAP) [138,139] is a visualization tool that can be used to make the output from the machine learning model more digestible. By computing the contribution of

each feature to the prediction, it can be used to explain the prediction of any model algorithm. To interpret the SHAP summary plot: the y-axis indicates the feature/predictors in order of importance from top to bottom; the value next to them is the mean SHAP value; the x-axis is the SHAP value; gradient color indicates the original value for the feature/predictor; finally, each point represents an observation from the original dataset.

4.8 ETHICAL CONSIDERATIONS

Research involving human subjects is guided by ethical principles to protect the participants from injury to their physical or mental health, or breach of their integrity. The World Medical Association stipulated what principles should be followed in the Helsinki Declaration issued in 1964, and it is generally agreed that scientific journals should not publish research that does not adhere to the Helsinki Declaration. While the document itself is not legally binding it has been codified in national legislation. In Sweden, principles of the Helsinki Declaration are included in the Ethical Review Act (2003:460), the Public Access and Secrecy Act (2009:300) and the Personal Data Act (1998:204), which was replaced by the EU directive the General Data Protection Regulation (GDPR) on May 25, 2018. The Ethical Review Act states that any research involving sensitive personal data, physical procedures in humans or biological samples from humans may only be performed with permission from the Swedish Ethical Review Authority (previously Regional Ethical Review Boards). The Ethical Review Authority weighs potential risk or harm to study participants against the expected benefits of new knowledge. As a general rule, study participants should be informed about the overarching study plan, aims, methods to be used, potential risks involved, who is responsible for the study, that participation is voluntary and that withdrawal from the study is allowed at any time. Once informed, participants must give consent to participation in order to be included in the study. When studies rely on secondary data only, such that no active participation is required by the study subjects, the Ethical Review Authority can waive the requirement of informed consent.

The data sources for this doctoral thesis were the population-based health registers in Sweden. Register holders will only release data intended for research if the research has been granted permission by the Ethical Review Authority. Based on the ethical approval, each register holder will make a secrecy assessment to determine what type of data, if any, can be released. The Regional Ethics Review Board in Stockholm granted ethical approval for the register linkage and aims included in this thesis in October 2013 (2013/1849-31), and for two amendments to expand the follow-up through a new linkage (2018/382-32) and formally shift the primary responsibility to the main supervisor (2019-01529). The recording of health information in registers may pose some risk for individuals, but Sweden permits use of such data for research under ethical permission. To minimize risks, data was pseudonymized, encrypted, and securely stored on dedicated servers at MEB, to which access is only granted to individuals directly involved in the management and analysis of the data. All handling, analysis, and storage of data was done in this secure server environment. Access requires proper login credentials with two-factor authentication and is logged and monitored.

5 RESULTS

5.1 STUDY I

The examination of the associations of ART with ADHD and school performance, and the role of underlying parental infertility, produced the following main results:

- In the cohort born between 1986 and 2001, 7% of births were from couples with infertility and 13% of these were achieved with ART, compared with 11% and 26%, respectively in the cohort of 1996 and 2012.
- The estimated cumulative incidence of ADHD by age 15 was 6.2% (95% CI: 5.9% to 6.5%) in children conceived with ART. This was lower in children born to couples with no known infertility (7.1%), and spontaneously conceived children of couples with infertility (7.3%, 95% CI: 7.2% to 7.5%) at the same age.
- Compared with all other children, children conceived with ART were at a lower risk of ADHD (Hazard ratio [HR]: 0.83; 95% CI:0.80 to 0.87). Adjustment for differences in parental characteristics flipped the association to a slightly higher risk in ART children (HR: 1.07, 95% CI: 1.02 to 1.12), although this was slightly attenuated by further adjustment for parental medical history (aHR: 1.05, 95% CI: 1.00 to 1.05). In the corresponding fully adjusted comparison to children of couples with infertility, however, no elevated risk was seen (aHR: 0.95, 95% CI: 0.90 to 1.00).
- The evaluation of school performance revealed similar patterns. Children conceived with ART had a higher average final grade (mean difference: 1.15, 95% CI: 1.09 to 1.21) and higher odds of being eligible for upper secondary school (Odds Ratio [OR]:1.53, 95% CI: 1.45 to 1.63) compared to all other children. Again, the associations were reversed with adjustment for parental characteristics and health history (mean difference in average grade: -0.14, 95% CI: -0.20 to -0.08; aOR for upper secondary school eligibility: 0.92, 95% CI: 0.85 to 0.96). The slight disadvantages were however fully attenuated in the comparison to children born to couples with infertility.
- No statistically significant differences were noted between specific ART procedures except that children conceived with ICSI were found at a slight advantage in school performance as compared with standard IVF.

5.2 STUDY II

The investigation of the associations of ART with mood disorders, anxiety, antidepressant use, and suicidal behavior yielded the following main results:

- Among all children born between 1994 and 2006, 10.4% were born to couples with infertility, of which 24.8% were conceived with ART. The median follow-up age was 18 (Interquartile range: 15-21).

- In the unexposed group of children born to couples with no known infertility, the estimated cumulative incidences of psychiatric health indices at the end of adolescence (18 years of age) were 5.6% for anxiety, 4.4% for a mood disorder, 7.0% for antidepressant use, and 0.9% for suicidal behavior, respectively. The corresponding cumulative incidences for children conceived with ART were overall comparable or lower. The remaining group of children born to couples with known infertility (but no use of ART) had a slightly higher cumulative incidence of anxiety (6.2%, 95% CI: 6.0% to 6.4%) and mood disorders (4.7%, 95% CI: 4.5% to 4.9%).
- Compared to all other, children conceived with ART were not at elevated risk of any of the outcomes except anxiety (HR: 1.07, 95% CI: 1.02 to 1.12), seemingly driven by OCD (HR: 1.35, 95% CI: 1.20 to 1.51). Adjustment for parental background characteristics had no influence on the association with any anxiety disorder (aHR: 1.08, 95% CI: 1.03 to 1.13) but attenuated that with OCD (aHR: 1.10, 95% CI: 0.98 to 1.24). No excess risk was observed for OCD or any other outcomes when the comparison was restricted to children born to couples with infertility.
- In the evaluation of specific ART procedures, children conceived with ICSI were not found any different from children of couples with infertility conceiving without ART, or with standard IVF respectively. Children conceived with fresh, but not frozen, embryo transfer had a lower risk of mood disorders compared with children of couples with infertility (aHR: 0.90, 95% CI: 0.83 to 0.97). In the direct contrast of the two procedures, frozen embryo transfer thus appeared associated with an elevated risk of mood disorders (aHR: 1.25, 95% CI: 1.03 to 1.52) and antidepressant use (aHR: 1.16, 95% CI: 1.01 to 1.33).

5.3 STUDY III

The assessment of ART and childhood asthma is part of an international collaboration to perform similar investigations in Sweden, Denmark, and Holland. Due to delays in the other countries, the result presentation is restricted to the findings from the Swedish study:

- Of all the children born between 2005 and 2013, 12% were from couples with infertility and 33% of those were conceived with ART. The estimated cumulative incidence of asthma at age 5 was 13.1% in children conceived with ART, slightly higher than the 12.7% in children born to couples with infertility, and 11.5% in all other children.
- Flexible parametric models suggested the hazards of asthma in ART and non-ART children were not proportional over time (attained age). In overall comparison (with all other children) children conceived with ART developed an excess risk of asthma after the first year of life, which then slightly declined around age 3 but remained elevated until the end of follow-up. With the full adjustment on the other hand, the hazard ratio was highest in the first year of life followed by a gradual decline toward the null around age 6. Finally, in the adjusted comparison with children of couples

with infertility, the attenuation was more rapidly such that the previously observed excess risk was largely limited to the first year of life.

- Examination of the association using stratified Cox regression with age bands (<1 year, 1-3 years, >3 years) revealed a similar pattern. While adjustment for parental characteristics brought out an elevated hazard ratio in the first year of life (aHR: 1.25, 95% CI: 1.18-1.33), the observed associations with ART after the first year were attenuated. In comparison with children of couples with infertility, the risk at age 1 remained elevated (aHR: 1.20, 95% CI: 1.11 to 1.29) while that of the other age bands was attenuated toward null. Finally, also the excess risk in the first year was substantially attenuated in the comparison to first maternal cousins (aHR: 1.10, 95% CI: 0.78-1.56).
- Children conceived via frozen embryo transfer were more likely to develop asthma during the first year of life only (aHR: 1.29, 95% CI: 1.13-1.50) in comparison to children conceived with fresh embryo transfer. However, ICSI did not seem to differ from standard IVF.
- Further comparing children conceived from specific ART procedures with children of couples with infertility, fresh embryo transfer showed no excess risk of asthma in the first year, whereas the other procedures did. As frozen embryo transfer had the most elevated risk (aHR: 1.40, 95% CI: 1.23-1.58), the different fertilization methods were reassessed among children conceived with fresh embryo transfer only, at which the previously noted excess risks were no longer seen.
- Compared to all other children, children conceived with ART were observed to have a higher risk of current asthma from age 3 through 6. The elevated risk was not seen in comparison with children of couples with infertility. No robust differences were found when further comparing different ART procedures.

5.4 STUDY IV

The attempts to develop prediction models for preeclampsia, placental complications, and PPH in nulliparous women undergoing ART yielded the following results:

- In total, 17,364 nulliparous women gave birth within the first three ART attempts between 2008 and 2016. Among these, 8,753 were from the first attempt, 5,397 from the second, and 3,214 from the third. The corresponding birth rates for each attempt were 24.4%, 23.0%, and 21.3% respectively. The couples were highly educated with more than 60% achieving post-secondary education. The women were on average 32 to 33 years old and the partners on average 3 years older.
- In those giving birth following their first attempt, ICSI accounted for over a third of the fertilizations, and cleavage embryo incubation dominated. As the attempts increased, so did the average gonadotropin dose, resulting in a slightly higher average number of oocytes retrieved. The use of ICSI also increased, as did the proportion of blastocyst embryo transfers as a result of frozen embryo transfer becoming an option in the subsequent attempts. Among women that gave birth following the first attempt,

5.4% had preeclampsia, 3.7% had placental complications, and 7.0% had PPH. The occurrence of preeclampsia and PPH increased with the number of attempts to achieve birth, but not for placental complications.

- In the pre-treatment prediction for those giving birth following the first attempt, model performances ranged from 55.4% to 61.8% for preeclampsia, 53.7% to 60.4% for placental complications, and 54.5% to 58.3% for PPH. No substantial improvements in the model predictions were seen in the post-treatment models or in the subsequent attempts. All model algorithms were performed similarly, and ensemble methods did not perform better than the simpler algorithms such as logistic regressions.
- In terms of the predictor importance, BMI, parental ages, and region of residence were the most influential demographic predictors, while infertility diagnosis was the most important medical history. In addition to the ART treatment information, year and stimulation protocol type were very influential, as was the type of embryo transfer (fresh or frozen) among women giving birth following subsequent attempts.

6 DISCUSSION

6.1 ART AND ADHD, MENTAL DISORDER, AND CHILDHOOD ASTHMA

6.1.1 The total effect of ART

Previous findings regarding links between ADHD and school performance have been conflicting. Early reports of normal attention and behavior from a small group of children conceived with ART in Holland [69,70] were followed by an alarmingly high prevalence of self-reported ADHD in a cross-sectional evaluation of young adults conceived with ART in the United States.[71] The Dutch team also assessed the school performance of 8-18 year-olds, finding no differences in relation to spontaneously conceived children from couples with infertility.[68] In contrast, children conceived with ART were found to have overall higher test scores compared with a random sample of spontaneously conceived children in a US study.[92] Results of our study align with findings from previously reported Scandinavian register-based studies, which found a lower risk of ADHD and slightly better overall school performance in children conceived with ART compared with all other children.[72-75]As in previous reports, the associations were reversed after adjustment for differences in parental characteristics.[72-75] However, when comparison was further restricted to children of couples with known infertility, no differences were seen, suggesting previously reported associations could have been due to residual confounding from parental factors related to the underlying infertility.

We expanded the investigation to assess the understudied mental health of adolescents conceived with ART, following a similar strategy to evaluate the role of parental background factors. First, children conceived with ART were at a slightly higher risk of anxiety and antidepressant use, seemingly driven by the elevated risk of OCD, whereas the risks of depression and suicidal behavior were lower. The observed elevated risks disappeared after accounting for parental background characteristics and underlying infertility. Before our study, the research on mental health of adolescents conceived with ART was scarce. Initial concerns came from an early study that relied on self-reports from parents and teachers [69] and the findings were not replicated in the follow-up query of the children themselves.[140] A register-based study from Finland scanned any mental disorders and reported a higher risk of any psychiatric diagnosis among singletons conceived with ART.[141] However, results for specific disorders such as depression and anxiety were underpowered.

In overall comparison to all other children, we also confirmed the previously reported link between parental ART use and asthma in children up to school-age.[77,79,81,142] However, we found the association with incident asthma to be most pronounced in the first year of life, and declining thereafter (non-proportional hazards with age). The risk of current asthma was found slightly elevated at age 3 through 6. In contrast to our findings for ADHD and school performance, adjustment for measured parental characteristics had a quite modest attenuating influence on the observed associations. When comparison was restricted to children with

parental infertility or to first-born maternal cousins, however, most of the association was attenuated to the null.

6.1.2 Specific ART procedures

In addition to investigating the total effect of ART on children's long-term outcomes we also evaluated the safety of specific optional ART procedures.

For the vast majority of outcomes under study, children conceived with ICSI were found no different from children conceived with standard IVF procedure. The only noted exception was a slightly higher probability of eligibility for upper secondary school. In light of the other findings and the relatively clear difference in indication (male factor) in the study period, this seems more likely due to residual confounding from parental factors, than an effect of the procedure per se. In sum, we found no indication for concern that this more invasive fertilization method would impact children's health with respect to the studied outcomes.

In a direct comparison of frozen and fresh embryo transfer, we found no difference with respect to ADHD or school performance. For psychiatric outcomes, the comparative safety assessment showed adolescents born after frozen embryo transfer were at elevated risk of mood disorder and antidepressant use. However, when compared to adolescents of couples with infertility (that conceived without ART) the frozen embryo transfer group was no different, whereas the fresh embryo transfer group had a lower risk of mood disorders. Finally, similar analysis in the evaluation of asthma suggested that the small remaining risk elevation in children conceived with ART seen only in the first year of life could be attributed to the group born from frozen embryo transfer.

Overall, our findings of specific ART procedures were mostly reassuring, where the only exception was the elevated risk of asthma in the first year of life following frozen embryo transfer. This finding aligns with a recent Swedish study that found a higher mortality of children conceived with frozen embryo transfer in the first year of life.[143] While the contrast between ICSI and standard IVF concerns only the fertilization method, the difference between frozen and fresh embryo transfer is not limited to the cryopreservation per se, but also the embryo stage at transfer and the maternal treatment protocol (instead of controlled ovarian hyperstimulation the transfer takes place in a natural, induced or inhibited ovulation cycle). Because of the mixed clinical practice in the implementation and previously inconsistent results, we find it difficult to explain our findings for frozen embryo transfer. Studies have shown singletons born after frozen embryo transfer are at lower risk of preterm birth, low birth weight, and small for gestational age [101,102] compared to children born after fresh embryo transfer. However, they also appear to be at elevated risk of being born large for gestational age and macrocosmic,[144,145] and the pregnancies are more likely complicated by preeclampsia and PPH.[102] The reasons behind these findings are not clearly understood but can include cryopreservation damages, influence of embryo transfer protocols, and epigenetic changes.[35,143,145-147] The original method of slow freezing has gradually been replaced by vitrification, and pre-treatment protocols that maintain the development of a corpus luteum

may be favored. Because of more favorable birth outcomes following frozen embryo transfers, elective freeze-all strategy, a method to freeze all the embryos after having eggs retrieved and implanted in later cycle(s), is increasingly being advocated for the avoidance of ovary hyperstimulation syndrome (OHSS), which only further stresses the need to understand potential negative effects of the approach. Some studies show improved results by applying a freeze-all policy but this strategy has both advantages and disadvantages and more research is needed to determine this issue, not least with respect to long-term health.[148,149] The fact that fresh and frozen embryo transfer is now being compared in randomized controlled trials [150,151] creates great potentials to do so in the coming years.

6.2 PREDICTING PREGNANCY COMPLICATIONS IN WOMEN UNDERGOING ART

Seeing whether patient and treatment characteristics could predict the risk of pregnancy complications in women undergoing ART, our models using information available at the start of treatment and antenatal care respectively did not achieve better prediction of preeclampsia, placental complications, and PPH than around 60%. The most commonly influential predictors across all outcomes were BMI, age, region of residence, infertility diagnosis, treatment year, stimulation protocol, and fresh or frozen embryo transfer.

Complex multi-factorial pathogenesis and unknown triggers have made preeclampsia very difficult to predict.[152] More recent prediction studies have included longitudinal clinical measurements such as repeated measures of blood pressure, ultrasound evaluations, and a variety of laboratory biomarkers, all of which have been shown to contribute to better prediction performance.[153] With regard to placental complications, there are to our knowledge no previous prediction model developments. For PPH, several known and influential risk factors, including mode of delivery and fetal size,[154] are not known until the time of labor and delivery, hampering prediction attempts in early pregnancy.

No matter whether it is related to the underlying infertility or in part due to the ART intervention, women undergoing ART are at a higher risk of complications in pregnancy and at delivery. Yet there are no prediction models specifically targeting this group. As these complications put both mother and baby at risk, and preventive/interventive measures are limited, the ultimate objective for prediction model development should be to provide personalized risk information to guide the subsequent pregnancy monitoring. To that goal, we considered a variety of predictors that would be readily available or easy to obtain during the course of ART treatment and at the time of enrolment to antenatal care respectively. However, since the suboptimal prediction models requires more work such as capturing more features and feature selection, we did not proceed to test the model on the holdout set (more in method consideration).

6.3 METHODOLOGICAL CONSIDERATION

6.3.1 Confounding by indication, and choice of comparison group

The overall findings in the studies of the thesis are consistently reassuring with respect to the health status of children conceived with ART. The confirmed associations from previous reports on these outcomes were largely attenuated after accounting for parental underlying infertility and background characteristics. The findings highlight the importance of considering confounding by indication in studies of ART safety. For example, mental disorders were more frequently seen in the children of couples with infertility.[87,88] Parental infertility has also been found associated with childhood asthma.[79] Children born to couples with infertility have been found a higher risk of ADHD.[87] No matter whether it is a common cause or more likely an effect of factors that in turn also influence outcomes in children, it is clear that infertility is an essential confounder, able to block critical confounding influence of the association between ART use and adverse outcomes (Figure 2.4).

The way to block such influence in the projects of the thesis is primarily the restriction of the comparison to children born to couples with infertility. This is because infertility is the predominant priori requirement for couples undergoing treatment (population at-risk). As such the restriction of couples with infertility reflects a more clinically relevant comparison and serves to balance parental background characteristics between exposed and unexposed. Couples with infertility that conceived without ART could have done so spontaneously or with the help of other types of fertility assistance, such as ovulation induction, intrauterine insemination, or other methods. Since our primary goal was to disentangle the effect of ART the restriction to couples with infertility was primarily aimed at decreasing the influence of unmeasured background characteristics related to the underlying infertility.

To maximize the ability to capture couples with infertility that conceived without ART, we considered both clinical diagnosis and self-reports. Using clinical diagnosis to identify infertility should result in few false positives (high specificity), but since not all couples with infertility may seek help or receive the diagnosis unless they also go on to receive treatment, this would miss many of the relevant cases for comparison (low sensitivity). It was therefore considered a strength to also be able to identify infertility from self-reported information in the MBR. The information is collected in the standardized interview by the midwife at the first antenatal visit, where women who report time to pregnancy of more than 1 year are also asked about potential fertility assistance. Again, specificity is expected to be high (few false positives), whereas we cannot exclude that some true cases could be missed due to failure to respond to, or correctly consider the question.

To account for the potential influence from shared stable family background factors, Study III also considered family-based designs to examine the association between ART and childhood asthma. Since twins cannot be discordant for ART, we primarily considered the use of sibling comparison to account for the suspected genetic influence on infertility and asthma in Study III, as this had been done in a previous study but with poor precision (power).[81] During

analysis, it became clear that this was not a straightforward approach. First of all, the reasons why siblings could be discordant for ART are very much dependent on other parental factors. On the one hand, nulliparous couples with trouble conceiving (primary infertility) often conceived their first child with ART. For various reasons, the underlying infertility may be resolved e.g., in oligo ovulation, or from stress reduction, enough that a sibling can be conceived without ART. On the other hand, as women age conception chances decrease, such that multiparous couples that were previously able to conceive without ART will need it for subsequent attempts (secondary infertility). When performing sibling comparison in either of these strata (ART in first born versus ART in later born), results varied dramatically so ignoring the differences would produce pooled results for sibling analysis that would depend on the ratio of primary and secondary infertility. Birth order appears to be associated with childhood asthma.[155] What's more the stratified analysis cannot separate a potential ART effect from the potential influence of birth order on the outcome. To avoid the influence of birth order while still achieving some control for familial factors, we finally considered comparisons of first-born maternal cousins, discordant for ART. Obviously, the level of confounder control was limited to factors that sisters share.

6.3.2 Other sources of confounding

Even though much confounding is addressed by the choice of a more appropriate comparison group, i.e., children of couples with infertility, there are still some background factors that may differ systematically between those conceived with ART and those that are not. For instance, some of the factors that are differently distributed among couples with infertility include parental ages, parental education, region of residence, and birth year. Parental ages, especially maternal age strongly influence both the type and severity of the underlying infertility, thus influencing the need for ART. While the majority of ART treatments and all other medical care in Sweden are publicly funded, parental education level influences health-seeking behavior as we found couples that conceived with ART more likely to have university-level education (>12 years) than all other couples with infertility. Heterogenous geographic distributions of medical services in Sweden have made the region of residence a proxy for access to specialized health care services, affecting both the likelihood of parental ART use as well as diagnosis and treatment of child outcomes. Birth year serves as a marker of some unmeasured changes over time, including the clinical practices of both IVF clinics and obstetrics and gynecology units, and diagnostic practices of the diseases of interest in the Swedish health care system.

6.3.3 Information bias

Information bias refers to measurement error or misclassification in continuous or categorical variables respectively, which distorts our estimation of the causal effect. In our studies, misclassification can occur in the measures for both the exposure and outcomes. If the misclassification of exposure differs with regard to the outcomes or vice versa, it refers to differential misclassification. Whereas differential misclassification can bias the causal effect

estimation in either direction (over or underestimation), non-differential misclassification tends to lead to an underestimation of the effect.

According to guidelines, it is mandatory for all IVF clinics in Sweden to report all assisted procedures and maintain good data quality. Still, the quality of the detailed assisted information varies among clinics, and on rare occasions, some clinics may miss to report due to administrative reasons. To account for potential missing reports, we also considered both self-reported use of ART from the first antenatal visit and diagnosis or procedure codes for ART treatments recorded in the NPR, under the assumption that the false positive rate of either of these sources should be low. The inclusion of the self-reports also allowed the detection of "ART treatment tourists" who traveled abroad for their ART treatments.

Identification of the outcomes of interest, such as ADHD, depression, and asthma should be reliable as the diagnoses are predominantly made by specialists such as child and adolescent psychiatrists and pediatricians, and children are also expected to be closely followed thereafter. However, we cannot rule out that some conditions/children can be diagnosed and treated in primary care. Therefore, we also considered medication use, alone as an alternative measure of psychiatric health and in combination with clinical diagnosis for asthma and ADHD based on validated algorithms to minimize the probability of misclassification in outcomes. Calendar year has an effect on misclassification as some data sources were established later (outpatient specialist visits since 2001, PDR since 2005). For the earliest born individuals, there is a risk that the incident event of interest was not captured, though to minimize this risk we considered different birth cohorts for the different outcomes under study (calibrated to the expected onset of disease).

In situations when investigating the effect of a disease or intervention as exposure on another disease as outcome, the identification of the outcome may not be solely decided by the natural course of the disease but by a heightened awareness due to the exposure. In such cases, exposed individuals may be more closely monitored, resulting in more diligent reporting or discovery of the outcomes in this group. This type of bias is referred to as detection bias. Especially in the early days of ART implementation, it seems likely that children would be subject to closer monitoring. As ART has been widely implemented and established as the primary method to overcome infertility, much of this closer monitoring is likely to have subsided, such that detection bias should not be a serious concern in our studies, still it cannot be excluded. Across all the studied outcomes, childhood asthma may be more vulnerable to detection bias as it can debut shortly after birth, when parents that had to undergo ART may still be more vigilant and/or anxious about the health of the child. At first look, this would fit with the fact that the only risk that we found not fully explained by parental characteristics was a small elevated risk of asthma in the first year of life. However, the excess risk appeared only among those conceived with frozen embryo transfer, and it is harder to see how detection bias would explain that, as we see no obvious reason for closer monitoring of children born from frozen embryo transfer compared with fresh embryo transfer.

In sum, any misclassification of ART should be non-differential as our ability to identify ART use had no relation to the outcomes. Such non-differential misclassification would lead to some underestimation of a potential causal association between ART and outcomes. In contrast, as just discussed, the misclassification of the outcome could possibly be differential in a manner that could lead to some overestimation due to detection bias.

6.3.4 Selection bias

Selection bias is when a non-causal association arises between exposure and outcome because we have conditioned on a common effect on a path between them. Selection bias is thus also synonymous with the previously described collider-stratification bias. In the studies of this thesis, we considered all live births in Sweden (singletons and multiples), performed multiple imputation for missing variables, and, with the exception of the assessment of school performance had no loss to follow-up, such that several sources of potential selection bias were eliminated. Moreover, most previous studies have assessed ART safety in singletons and twins separately (or most commonly singletons only). Since multiple gestations is a consequence of the intervention, such practice has the potential to open a backdoor path between ART and later health outcomes if these share risk factors with multiple gestation (which then becomes a collider; figure 2.4). Therefore, we chose not to stratify our studies with respect to multiple gestation.

6.3.5 Generalizability

Generalizability refers to the degree the findings of a study are transferable to other settings and populations. The generalizability, or external validity, of the findings in the thesis may be limited to the setting in which the studies took place, i.e., the time period and clinical practice in Sweden. In particular, there have been several considerable changes to clinical practice during the study period, including: 1) Sweden's implementation of single embryo transfer in the early 2000s; 2) use of slow freezing methods being gradually replaced by vitrification; 3) ICSI being used more broadly, i.e., not solely for male factor infertility. Aside from the use of ICSI these changes are broadly seen as improvements, though the vitrification will require follow-up in future studies.

6.3.6 Method consideration for Study IV

Overfitting refers to a situation when a model fits too well on training data. This has been a major challenge in machine learning. Among many, the test data is not seen at the stage of model development, and hyperparameter optimization can overfit a dataset and provide an optimistic evaluation of a model that should not be used for model selection.[156] Cross-validation is one of the easy methods to minimize the probability of overfitting. However, in the comprehensive modeling approach of Study IV we attempted to both compare different algorithms and tuning hyperparameters. To minimize the risk of overfitting, we therefore considered nested cross-validation method.

Another consideration is feature selection, which refers to the method in the selection of predictors or variables. Feature selection includes consideration of input predictors but mostly refers to the process of reducing the number of input variables in the model constructions. When the goal is to obtain a prediction model that requires limited data collection, as is typically the case for daily clinical practice, feature selection should be applied to identify the most parsimonious model. Methods used in feature selection involve both supervised learning and unsupervised learning such as principal component analysis (PCA), clustering, pattern recognition, etc., but also selection informed by domain knowledge or feature importance. Despite the inclusion of demographic and lifestyle factors, medical history, and details about the ART treatment, the prediction performances were not satisfying. While some improvement may still be made from the inclusion of additional background information (e.g., miscarriage history), more work on feature categorization (e.g., dealing with correlation between features), and reducing the dimension of features, the critical step to achieving the ultimate goal of clinical prediction will be to build longitudinal prediction models incorporating repeated pregnancy measures.

Since the overall prediction performances were suboptimal, we did not continue to test the model on the holdout set, rather saving it for the aforementioned fine-tuning of the feature selection. However, the extensive implementation and comparison of different methods provide some reassurance that the prediction of popular algorithms in machine learning is not necessarily better performing than standard logistic regression. Furthermore, the feature importance evaluation from our models suggests that patient and treatment characteristics may be relevant predictors of pregnancy complications in this patient group, in particular such that is related to the maternal hormonal profile and whether the embryo transfer was fresh or frozen.

7 CONCLUSIONS

This nationwide long-term follow-up of children conceived with ART found overall favorable outcomes with respect to school performance, risk of ADHD and psychiatric health compared to all other children/adolescents. Observed elevated risks of OCD and asthma in infancy were largely explained by differences in parental characteristics including parental underlying infertility, which could be a candidate for potential screening. The findings provide broad reassurance concerning offspring health and development following use of ART, and stress the importance of considering the role of underlying indication (and its associated risk factors) in ART safety studies. In the evaluation of safety, null findings are welcome, and we are pleased to be able to provide reassurance to ART patients and providers. Our finding of elevated risk of asthma in the first year of life among children born after frozen embryo transfer warrants further confirmation and scrutiny in future studies.

Predictions of complications during pregnancy and delivery in women undergoing ART treatment were also attempted, however, prediction performances were suboptimal. The most commonly influential predictors were BMI, parental age, treatment year, region of residence, infertility diagnosis, stimulation protocol, and fresh/frozen embryo transfer.

8 POINTS OF PERSPECTIVE

ART has been widely accepted and implemented worldwide. Understanding the long-term safety of ART is increasingly relevant and important as a growing population of children conceived with ART are going through adolescence and reaching adulthood.

Further research with long-term follow-up of large samples and rigorous methods is necessary to thoroughly evaluate the safety of ART. Nationwide register-based studies will continue to be vital to the assessment of long-term safety, offering complete population coverage of reliable information on ART use and different clinical endpoints.

Future studies must continue to consider and properly account for the role of underlying parental characteristics (selected due to infertility). If an association remains, further investigation should consider separating the influence of ART procedures from downstream factors on the causal pathways. Finally, the comparative safety of newer, optional, and increasingly commonly implemented specific procedures needs to be evaluated. This includes disentangling the effect of ICSI use from its (earlier) indication and understanding the long-term effects of cryopreservation methods and implantation protocols.

Predicting pregnancy complications following ART may not be possible at the start of treatment and/or antenatal care, but future attempts incorporating longitudinal clinical measures in pregnancy should consider patient and treatment characteristics in the prediction among this group of women.

9 ACKNOWLEDGEMENTS

Time flies. Leaving far away from home, settling down in a country with a completely different culture, and having been working remotely during the period could have made my doctoral life so hard if I would not be helped by people with beautiful hearts along the way. I hereby sincerely thank the following people with my heartfelt gratitude:

Anna Sara Oberg: my whole-hearted main supervisor and friend. When I am drafting this section, all of those memorable scenes between us came back to me like a movie. Our journey started from the day we first met in the interview where you said the doctoral ship is like a marriage. There is absolutely no doubt that choosing the right supervisor is equally important as choosing the research field, and it is my great honor to be your very first Ph.D. student. Sincerely thank you for the countless time you have spent and been patiently teaching me epidemiological methods (all these whole afternoon discussions, overtime zoom personal meetings). Although the most of time you were not physically around me, you were always available for taking any question from me, which is much more important. Thank you for having been always supportive and encouraging of my writing. What I have learned from you is far more than knowledge, and I feel so fortunate to have you as my supervisor and friend in the past, and also in the future.

Sonia Hernández-Díaz: my co-supervisor. Perhaps you never know that both you and Sara have always been the role model to me and show me what a scientist should be like - to be modest and open-minded. I enjoyed every second of the inspiring and in-depth discussions between you and Sara. Additionally, thank you for inviting me to visit Harvard T.H. Chan School of Public Health where I learned by sneaking into many classes and seminars and got inspired by sages and the talented.

Catarina Almqvist Malmros: my co-supervisor. You have always been so kind and thoughtful. Thank you for offering me the opportunity to participate and present in your research group meetings. Borrowing from both you and the groupmates, I easily obtained many clinical and methodological considerations and invaluable input for both childhood asthma and family design methods. Thank you very much for the memorable Christmas host and lottery gift (sourdough)!

Anna L. V. Johansson: my co-supervisor and biostatistician. Thank you for teaching me some of these essential statistical models, especially survival models, and guiding me on what to explore statistically for a deeper understanding of the findings. Moreover, thank you for the lovely Lucia songs these years that always delight us in the dark winter.

Sandra Eloranta: my dear mentor, for all of the lovely conversations and coffee times (Fika). Thank you for finding and helping me when I struggled with a great amount of uncertainty. I am bowled over by those of your magnificent travels in China, and your passionate and endless energies on the things you like. More importantly, thank you for always being available listening to me when I sometimes found my life in confusion, and for

guiding me to find what I should pursue thereafter. Your strong self-discipline has made me admire but also ashamed. I feel so lucky to meet you in my life and feel so grateful to have you as my mentor!

Marie Reilly: my former supervisor in my master thesis. I still remembered the task, the corresponding effort, and the following interview I had made to be qualified to do a master thesis in your group. I was so nervous until you said my assignment was a masterpiece. Thank you for introducing a new world (register-based studies) to me, and offering me the opportunity to learn data management, and statistical programming, all of which have a profound impact on what I am now pursuing and my future career. You are also a lovely and quick-witted person. Your tricky and sharp questions, critical comments, friendly advice, and patient explanations drive me forward from different perspectives. All of these fun discussions and explorations in the database made my master-student period super enjoyable and memorable.

Cina Nyberg: my both senior and junior fellow Ph.D. student. Thank you for your solid clinical knowledge of infertility and ART, and also for offering me the opportunity for a very unique experience in the IVF clinic and laboratory as well as meeting patients. You also show me how research and learning can be driven by pure interests and it is never too late.

To my other co-authors **Mikael Landén, Kenny Rodriguez-Wallberg,** and **Anuj Pareek,** sincerely thank you for generously sharing clinical or methodological knowledge and valuable input.

Thank you, **Mariam Lashkariani,** and **Staffan Bergh**, for the great work and help on data administration and management. Thank you, **Alessandra Nanni,** for the study administration during my Ph.D. period. Thank you, **Vivekananda Lanka**, for patiently teaching me SAS programming and having been there available to help me troubleshoot when I was a junior.

Thank you doctor **Sipan Khalil,** psychologist **Lina Pleijel,** and physiotherapist **Jenny Larsson** for the professional help when I felt my life surrounded by panic, anxiety, and negative thoughts.

Malin Rytterdahl and **Jonathan Björkman** couple, my Swedish lovely friends, thank you for celebrating my birthday, and those joyful and memorable treats for the overnight stays, wish you the brightest career, Jona, and looking forward to staying at the house of your own land next time, once you have it! **Gabriele Girelli**, you are the smartest person I have seen at our age. Looking forward to having more fun occasions and events in the future including but limited to cultures of East Asia, food, movies, and games. **Sam**, the dearest Swedish friend of my master student's period, I lost contact with you since the end of my master thesis. You have the most beautiful heart. Now I have almost finished my Ph.D. and wonder how you have been during these years. I hope all of these noises, and negative thoughts will no longer chase you and wish you a much brighter life ahead.

Donghao Lu, my undergraduate university (Sichuan University) alumni and senior fellow. Being like a big brother, thank you for these informative conversations and food occasions not only in Sweden but also in Boston. Meeting you in Boston made me feel at home. **Zheng Ning**, my closest friend during doctoral period, Thank you for the invaluable discussions of methods and countless companies along the way. Our friendship started at office 5134 and will continue for life. Looking forward to keeping chatting with you about games, dieting, medical science, statistics, and hiking (the order is completely random). Indeed, we have a lot to learn from each other. 一天不学习，浑身难受。Also, many thanks to **Yinxi Wang**, and **Yunzhang Wang**, for the invaluable and inspired discussions on Study IV.

I would also like to thank the rest of my Chinese friends here in Sweden: **Cen Chen, Yuxin Cheng and Chang Zhao** couple**, Xiaoyang Du, Shuang Hao, Huanling Lai and Kevin Kamimura** couple**, Xinhe Mao, Lili Meng, Jie Song, Ximeng Su, Rongrong Wu, Mujin Ye, Erwei Zeng and Hengsha Li** couple**, Le Zhang, Yuqi Zhang, Yitian Zhou, and Family Pang Isaksson.** 越长大越孤单，衷心的感谢所有的朋友能够包容一个与世界不太相同的我。在这些孤单又平凡的日子里，在这世界上的遥远的角落里，有你们真好。

Aven He, my best and most like-minded friend in my life. Thank you for your always available remote consultation from China, no matter when. 何元周天，睡在我大学寝室上铺的四字朋友，我的拜把子兄弟，一晃又是十年，何时出来转转呀。期待在未来的后疫情时代里，与你重逢。

At last, I would like to sincerely thank **my family** and my beloved other half: **Xinge Li**. 衷心的感谢生我养我的父母，特别是不善言辞的老爸。爸你说过，好男儿志在四方，结果我就头也不回地跑这么远，感谢你们让我健康的成长。感谢冬妮的爸妈，信任地把宝贝女儿托付给我。欣格，感谢让我在这世界的角落遇到你，我们的生活存在于柴米油盐，也在繁星下的阿尔卑斯山。你是我博士寒夜中的一束光，有你和豆豆在的地方，便是家！

在这患上焦虑和社恐的博士最后的日子里，我愈发的珍惜身边的人们，真心祝愿每一个爱我和帮助过我的朋友，万事顺遂，锦绣前程。

10 REFERENCES

1. Zegers-Hochschild F, Adamson GD, de Mouzon J, et al. International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. *Fertility and sterility : official of the American society for reproductive medicine*. 2009;92(5):1520-1524. doi:10.1016/j.fertnstert.2009.09.009

2. Gnoth C, Godehardt D, Godehardt E, Frank-Herrmann P, Freundl G. Time to pregnancy: results of the German prospective study and impact on the management of infertility. *Human reproduction*. 2003;18(9):1959-1966. doi:10.1093/humrep/deg366

3. Van Voorhis BJ. Clinical practice. In vitro fertilization. *The New England journal of medicine*. 2007;356(4):379-386. doi:10.1056/NEJMcp065743

4. Audibert C, Glass D. A global perspective on assisted reproductive technology fertility treatment: an 8-country fertility specialist survey. *Reproductive biology and endocrinology : RB & E*. 2015;13:133. doi:10.1186/s12958-015-0131-z

5. Wyns C, Bergh C, Calhaz-Jorge C, et al. ART in Europe, 2016: results generated from European registries by ESHRE. *Hum Reprod Open*. 2020;2020(3):hoaa032. doi:10.1093/hropen/hoaa032

6. Effectiveness and treatment for unexplained infertility. *Fertility and sterility : official of the American society for reproductive medicine*. 2006;86(5 Suppl 1):S111-S114. doi:10.1016/j.fertnstert.2006.07.1475

7. Healy DL, Trounson AO, Andersen AN. Female infertility: causes and treatment. *The Lancet*. 1994;343(8912):1539-1544. doi:10.1016/s0140-6736(94)92941-6

8. Fertility: Assessment and Treatment for People with Fertility Problems. (Royal College of Obstetricians & Gynaecologists National Collaborating Centre for Women's and Children's Health) (2013).

9. Ata B, Seyhan A, Seli E. Diminished ovarian reserve versus ovarian aging: overlaps and differences. *Current opinion in obstetrics and gynecology*. 2019;31(3):139-147. doi:10.1097/GCO.0000000000000536

10. Dunson DB, Baird DD, Colombo B. Increased infertility with age in men and women. *Obstetrics and gynecology*. 2004;103(1):51-56. doi:10.1097/01.AOG.0000100153.24061.45

11. Magnus MC, Wilcox AJ, Morken N-H, Weinberg CR, Håberg SE. Role of maternal age and pregnancy history in risk of miscarriage: prospective register based study. *BMJ : British medical journal*. 2019;364:l869. doi:10.1136/bmj.l869

12. Matthews TJ, Hamilton BE. First births to older women continue to rise. *NCHS Data Brief*. 2014;(152):1-8.

13. Bureau USC. Estimated Median Age at First Marriage, by Sex: 1890 to the Present. In: Table MS-2. Estimated Median Age at First Marriage bSttP, editor. The United States: The U.S. Census Bureau's Public Information Office; 2019.

14. Schoen R, Canudas-Romo V. Timing effects on first marriage: twentieth-century experience in England and Wales and the USA. *Population studies*. 2005;59(2):135-146. doi:10.1080/00324720500099124

15. Lunenfeld B. Historical perspectives in gonadotrophin therapy. *Hum Reprod Update*. 2004 Nov-Dec 2004;10(6):453-67. doi:10.1093/humupd/dmh044

16. Dorn C, van der Ven H. Clomiphene citrate versus gonadotrophins for ovulation stimulation. *Reprod Biomed Online*. Jun 2005;10 Suppl 3:37-43. doi:10.1016/s1472-6483(11)60389-1

17. Biggers JD. In vitro fertilization and embryo transfer in human beings. *The New England journal of medicine (print)*. 1981;304(6):336-342. doi:10.1056/NEJM198102053040607

18. Steptoe PC, Edwards RG. Birth after the reimplantation of a human embryo. *The Lancet*. 1978;2(8085):366. doi:10.1016/s0140-6736(78)92957-4

19. Elias S, Annas GJ. Social policy considerations in noncoital reproduction. *JAMA : the journal of the American Medical Association*. 1986;255(1):62-68.

20. Wagner MG, St Clair PA. Are in-vitro fertilisation and embryo transfer of benefit to all? *The Lancet*. 1989;2(8670):1027-1030. doi:10.1016/s0140-6736(89)91029-5

21. Palermo G, Joris H, Devroey P, Van Steirteghem AC. Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. *The Lancet*. 1992;340(8810):17-18. doi:10.1016/0140-6736(92)92425-f

22. Dyer S, Chambers GM, de Mouzon J, et al. International Committee for Monitoring Assisted Reproductive Technologies world report: Assisted Reproductive Technology 2008, 2009 and 2010. *Human reproduction*. 2016;31(7):1588-1609. doi:10.1093/humrep/dew082

23. Rodriguez-Wallberg KA, Waterstone M, Anastácio A. Ice age: Cryopreservation in assisted reproduction - An update. *Reprod Biol*. Jun 2019;19(2):119-126. doi:10.1016/j.repbio.2019.04.002

24. Edgar DH, Gook DA. A critical appraisal of cryopreservation (slow cooling versus vitrification) of human oocytes and embryos. *Hum Reprod Update*. 2012 Sep-Oct 2012;18(5):536-54. doi:10.1093/humupd/dms016

25. Rienzi L, Gracia C, Maggiulli R, et al. Oocyte, embryo and blastocyst cryopreservation in ART: systematic review and meta-analysis comparing slow-freezing versus vitrification to produce evidence for the development of global guidance. *Hum Reprod Update*. 03 01 2017;23(2):139-155. doi:10.1093/humupd/dmw038

26. Centers for Disease Control and Prevention, American Society for Reproductive Medicine, Society for Assisted Reproductive Technology. *2016 Assisted Reproductive Technology Fertility Clinic and National Summary Report.* 2018.

27. Swedish Society of Obstetrics and Gynecology (SFOG). Reproduktionsmedincin. Swedish Society of Obstetrics and Gynecology; 2019.

28. Nationellt kvalitetsregister för assisterad befruktning (Q-IVF). Fertility treatments in Sweden--National report 2021. 2021.

29. Steegers EAP, von Dadelszen P, Pijnenborg R, et al. Pre-eclampsia. *The Lancet*. 2010;376(9741):631-644. doi:10.1016/S0140-6736(10)60279-6

30. Fujitake E, Jaspal R, Monasta L, Stampalija T, Lees C. Acute cardiovascular changes in women undergoing in vitro fertilisation (IVF), a systematic review and meta-analysis. *Eur J Obstet Gynecol Reprod Biol*. May 2020;248:245-251. doi:10.1016/j.ejogrb.2020.01.033

31. Wang YA, Chughtai AA, Farquhar CM, Pollock W, Lui K, Sullivan EA. Increased incidence of gestational hypertension and preeclampsia after assisted reproductive technology treatment. *Fertil Steril*. Apr 2016;105(4):920-926.e2. doi:10.1016/j.fertnstert.2015.12.024

32. Chih HJ, Elias FTS, Gaudet L, Velez MP. Assisted reproductive technology and hypertensive disorders of pregnancy: systematic review and meta-analyses. *BMC Pregnancy Childbirth*. Jun 28 2021;21(1):449. doi:10.1186/s12884-021-03938-8

33. Martin AS, Monsour M, Kawwass JF, Boulet SL, Kissin DM, Jamieson DJ. Risk of Preeclampsia in Pregnancies After Assisted Reproductive Technology and Ovarian Stimulation. *Matern Child Health J*. 10 2016;20(10):2050-6. doi:10.1007/s10995-016-2067-0

34. Sha T, Yin X, Cheng W, Massey IY. Pregnancy-related complications and perinatal outcomes resulting from transfer of cryopreserved versus fresh embryos in vitro fertilization: a meta-analysis. *Fertility and sterility*. 2018;109(2):330-342.e9. doi:10.1016/j.fertnstert.2017.10.019

35. Moreno-Sepulveda J, Espinós JJ, Checa MA. Lower risk of adverse perinatal outcomes in natural versus artificial frozen-thawed embryo transfer cycles: a systematic review and meta-analysis. *Reproductive biomedicine online*. 2021;42(6):1131-1145. doi:10.1016/j.rbmo.2021.03.002

36. Singh B, Reschke L, Segars J, Baker VL. Frozen-thawed embryo transfer: the potential importance of the corpus luteum in preventing obstetrical complications. *Fertil Steril*. 02 2020;113(2):252-257. doi:10.1016/j.fertnstert.2019.12.007

37. Oberg AS, VanderWeele TJ, Almqvist C, Hernandez-Diaz S. Pregnancy complications following fertility treatment-disentangling the role of multiple gestation. *International journal of epidemiology*. 2018;47(4):1333-1342. doi:10.1093/ije/dyy103

38. Vermey BG, Buchanan A, Chambers GM, et al. Are singleton pregnancies after assisted reproduction technology (ART) associated with a higher risk of placental anomalies compared with non-ART singleton pregnancies? A systematic review and meta-analysis. *BJOG*. Jan 2019;126(2):209-218. doi:10.1111/1471-0528.15227

39. Tai W, Hu L, Wen J. Maternal and Neonatal Outcomes After Assisted Reproductive Technology: A Retrospective Cohort Study in China. *Front Med (Lausanne)*. 2022;9:837762. doi:10.3389/fmed.2022.837762

40. Sacha CR, Harris AL, James K, et al. Placental pathology in live births conceived with in vitro fertilization after fresh and frozen embryo transfer. *Am J Obstet Gynecol*. 04 2020;222(4):360.e1-360.e16. doi:10.1016/j.ajog.2019.09.047

41. Cirillo F, Grilli L, Ronchetti C, et al. Retrospective comparison of pregnancy outcomes of fresh and frozen-warmed single blastocyst transfer: a 5-year single-center experience. *J Assist Reprod Genet*. Jan 2022;39(1):201-209. doi:10.1007/s10815-021-02362-3

42. Martin PM, Welch HG. Probabilities for singleton and multiple pregnancies after in vitro fertilization. *Fertility and sterility : official of the American society for reproductive medicine*. 1998;70(3):478-481. doi:10.1016/s0015-0282(98)00220-9

43. Pandian Z, Templeton A, Serour G, Bhattacharya S. Number of embryos for transfer after IVF and ICSI: a Cochrane review. *Human reproduction*. 2005;20(10):2681-2687. doi:10.1093/humrep/dei153

44. Busnelli A, Dallagiovanna C, Reschini M, Paffoni A, Fedele L, Somigliana E. Risk factors for monozygotic twinning after in vitro fertilization: a systematic review and meta-analysis. *Fertility and sterility*. 2019;111(2):302-317. doi:10.1016/j.fertnstert.2018.10.025

45. Chang HJ, Lee JR, Jee BC, Suh CS, Kim SH. Impact of blastocyst transfer on offspring sex ratio and the monozygotic twinning rate: a systematic review and meta-analysis. *Fertility and sterility*. 2009;91(6):2381-2390. doi:10.1016/j.fertnstert.2008.03.066

46. Ferraretti AP, Nygren K, Andersen AN, et al. Trends over 15 years in ART in Europe: an analysis of 6 million cycles. *Human reproduction open*. 2017;2017(2):hox012. doi:10.1093/hropen/hox012

47. Wright VC, Schieve LA, Reynolds MA, Jeng G, Kissin D. Assisted reproductive technology surveillance -- United States, 2001. *Morbidity and mortality weekly report MMWR*. 2004;53:1-20.

48. Prevention of twin pregnancies after IVF/ICSI by single embryo transfer. ESHRE Campus Course Report. *Human reproduction*. 2001;16(4):790-800.

49. Jackson RA, Gibson KA, Wu YW, Croughan MS. Perinatal outcomes in singletons following in vitro fertilization: a meta-analysis. *Obstetrics and gynecology*. 2004;103(3):551-563. doi:10.1097/01.AOG.0000114989.84822.51

50. Helmerhorst FM, Perquin DAM, Donker D, Keirse MJNC. Perinatal outcome of singletons and twins after assisted conception: a systematic review of controlled studies. *BMJ : British medical journal*. 2004;328(7434):261. doi:10.1136/bmj.37957.560278.EE

51. Hansen M, Bower C, Milne E, de Klerk N, Kurinczuk JJ. Assisted reproductive technologies and the risk of birth defects--a systematic review. *Human reproduction*. 2005;20(2):328-338. doi:10.1093/humrep/deh593

52. Hyrapetian M, Loucaides EM, Sutcliffe AG. Health and disease in children born after assistive reproductive therapies (ART. *Journal of reproductive immunology*. 2014;106:21-26. doi:10.1016/j.jri.2014.08.001

53. Hansen M, Kurinczuk JJ, Milne E, de Klerk N, Bower C. Assisted reproductive technology and birth defects: a systematic review and meta-analysis. *Human reproduction update*. 2013;19(4):330-353. doi:10.1093/humupd/dmt006

54. Ghazi HA, Spielberger C, Källén B. Delivery outcome after infertility--a registry study. *Fertility and sterility : official of the American society for reproductive medicine*. 1991;55(4):726-732. doi:10.1016/s0015-0282(16)54238-1

55. Schreurs A, Legius E, Meuleman C, Fryns JP, D'Hooghe TM. Increased frequency of chromosomal abnormalities in female partners of couples undergoing in vitro fertilization or intracytoplasmic sperm injection. *Fertility and sterility : official of the American society for reproductive medicine*. 2000;74(1):94-96. doi:10.1016/s0015-0282(00)00558-6

56. Rimm AA, Katayama AC, Katayama KP. A meta-analysis of the impact of IVF and ICSI on major malformations after adjusting for the effect of subfertility. *Journal of assisted reproduction and genetics*. 2011;28(8):699-705. doi:10.1007/s10815-011-9583-z

57. Lazaraviciute G, Kauser M, Bhattacharya S, Haggarty P, Bhattacharya S. A systematic review and meta-analysis of DNA methylation levels and imprinting disorders in

children conceived by IVF/ICSI compared with children conceived spontaneously. *Human reproduction update*. 2014;20(6):840-852. doi:10.1093/humupd/dmu033

58. Sato A, Otsu E, Negishi H, Utsunomiya T, Arima T. Aberrant DNA methylation of imprinted loci in superovulated oocytes. *Human reproduction*. 2007;22(1):26-35. doi:10.1093/humrep/del316

59. Hammoud SS, Purwar J, Pflueger C, Cairns BR, Carrell DT. Alterations in sperm DNA methylation patterns at imprinted loci in two classes of infertility. *Fertility and sterility : official of the American society for reproductive medicine*. 2010;94(5):1728-1733. doi:10.1016/j.fertnstert.2009.09.010

60. Batcheller A, Cardozo E, Maguire M, DeCherney AH, Segars JH. Are there subtle genome-wide epigenetic alterations in normal offspring conceived by assisted reproductive technologies? *Fertility and sterility : official of the American society for reproductive medicine*. 2011;96(6):1306-1311. doi:10.1016/j.fertnstert.2011.09.037

61. Middelburg KJ, Heineman MJ, Bos AF, Hadders-Algra M. Neuromotor, cognitive, language and behavioural outcome in children born following IVF or ICSI-a systematic review. *Human reproduction update*. 2008;14(3):219-231. doi:10.1093/humupd/dmn005

62. Bay B, Mortensen EL, Kesmodel US. Assisted reproduction and child neurodevelopmental outcomes: a systematic review. *Fertility and sterility*. 2013;100(3):844-853. doi:10.1016/j.fertnstert.2013.05.034

63. Hart R, Norman RJ. The longer-term health outcomes for children born as a result of IVF treatment. Part II-Mental health and development outcomes. *Human reproduction update*. 2013;19(3):244-250. doi:10.1093/humupd/dmt002

64. Klemetti R, Sevón T, Gissler M, Hemminki E. Health of children born as a result of in vitro fertilization. *Pediatrics /*. 2006;118(5):1819-1827. doi:10.1542/peds.2006-0735

65. Maimburg RD, Vaeth M. Do children born after assisted conception have less risk of developing infantile autism? *Human reproduction*. 2007;22(7):1841-1843. doi:10.1093/humrep/dem082

66. Mazzotti S, Calderoni S, Saviozzi I, Guzzetta A, Conti E. Are children born after assisted reproductive technology at increased risk of autism spectrum disorders? A systematic review. *Human reproduction*. 2013;28(12):3316-3327. doi:10.1093/humrep/det380

67. Sandin S, Nygren K-G, Iliadou A, Hultman CM, Reichenberg A. Autism and Mental Retardation Among Offspring Born After In Vitro Fertilization. *JAMA : the journal of the American Medical Association*. 2013;310(1):75-84. doi:10.1001/jama.2013.7222

info:doi/10.1001/jama.2013.7222

68. Wagenaar K, Ceelen M, van Weissenbruch MM, Knol DL, Delemarre-van de Waal HA, Huisman J. School functioning in 8- to 18-year-old children born after in vitro fertilization. *European journal of pediatrics*. 2008;167(11):1289-1295. doi:10.1007/s00431-008-0677-2

69. Wagenaar K, van Weissenbruch MM, Knol DL, Cohen-Kettenis PT, Delemarre-van de Waal HA, Huisman J. Behavior and socioemotional functioning in 9-18-year-old children born after in vitro fertilization. *Fertility and sterility*. 2009;92(6):1907-1914. doi:10.1016/j.fertnstert.2008.09.026

70. Wagenaar K, van Weissenbruch MM, Knol DL, Cohen-Kettenis PT, Delemarre-van de Waal HA, Huisman J. Information processing, attention and visual-motor function of adolescents born after in vitro fertilization compared with spontaneous conception. *Human reproduction*. 2009;24(4):913-921. doi:10.1093/humrep/den455

71. Beydoun HA, Sicignano N, Beydoun MA, et al. A cross-sectional evaluation of the first cohort of young adults conceived by in vitro fertilization in the United States. *Fertility and sterility*. 2010;94(6):2043-2049. doi:10.1016/j.fertnstert.2009.12.023

72. Bay B, Mortensen EL, Hvidtjørn D, Kesmodel US. Fertility treatment and risk of childhood and adolescent mental disorders: register based cohort study. *BMJ : British medical journal*. 2013;347:f3978. doi:10.1136/bmj.f3978

73. Källén AJB, Finnström OO, Lindam AP, Nilsson EME, Nygren K-G, Otterblad Olausson PM. Is there an increased risk for drug treated attention deficit/hyperactivity disorder in children born after in vitro fertilization? *European journal of paediatric neurology*. 2011;15(3):247-253. doi:10.1016/j.ejpn.2010.12.004

74. Norrman E, Petzold M, Bergh C, Wennerholm U-B. School performance in singletons born after assisted reproductive technology. *Human reproduction*. 2018;33(10):1948-1959. doi:10.1093/humrep/dey273

75. Spangmose AL, Malchau SS, Schmidt L, et al. Academic performance in adolescents born after ART-a nationwide registry-based cohort study. *Human reproduction*. 2017;32(2):447-456. doi:10.1093/humrep/dew334

76. Bjerg A, Sandström T, Lundbäck B, Rönmark E. Time trends in asthma and wheeze in Swedish children 1996-2006: prevalence and risk factors by sex. *Allergy*. Jan 2010;65(1):48-55. doi:10.1111/j.1398-9995.2009.02105.x

77. Carson C, Sacker A, Kelly Y, Redshaw M, Kurinczuk JJ, Quigley MA. Asthma in children born after infertility treatment: findings from the UK Millennium Cohort Study. *Human reproduction*. 2013;28(2):471-479. doi:10.1093/humrep/des398

78. Guibas GV, Moschonis G, Xepapadaki P, et al. Conception via in vitro fertilization and delivery by Caesarean section are associated with paediatric asthma incidence. *Clinical & experimental allergy*. 2013;43(9):1058-1066. doi:10.1111/cea.12152

79. Källén B, Finnström O, Nygren K-G, Otterblad Olausson P. Asthma in Swedish children conceived by in vitro fertilisation. *Archives of disease in childhood*. 2013;98(2):92-96. doi:10.1136/archdischild-2012-301822

80. Kuiper DB, Seggers J, Schendelaar P, et al. Asthma and asthma medication use among 4-year-old offspring of subfertile couples--association with IVF? *Reproductive biomedicine online*. 2015;31(5):711-714. doi:10.1016/j.rbmo.2015.08.002

81. Magnus MC, Karlstad Ø, Parr CL, et al. Maternal history of miscarriages and measures of fertility in relation to childhood asthma. *Thorax*. 2019;74(2):106-113. doi:10.1136/thoraxjnl-2018-211886

82. Finnström O, Källén B, Lindam A, Nilsson E, Nygren K-G, Olausson PO. Maternal and child outcome after in vitro fertilization--a review of 25 years of population-based data from Sweden. *Acta obstetricia et gynecologica Scandinavica*. 2011;90(5):494-500. doi:10.1111/j.1600-0412.2011.01088.x

83. Ericson A, Nygren KG, Olausson PO, Källén B. Hospital care utilization of infants born after IVF. *Human reproduction*. 2002;17(4):929-932. doi:10.1093/humrep/17.4.929

84. Baird DD, Wilcox AJ, Kramer MS. Why might infertile couples have problem pregnancies? *The Lancet*. 1999;353(9166):1724-1725. doi:10.1016/S0140-6736(99)90050-8

85. Basso O, Baird DD. Infertility and preterm delivery, birthweight, and Caesarean section: a study within the Danish National Birth Cohort. *Human reproduction*. 2003;18(11):2478-2484. doi:10.1093/humrep/deg444

86. Källén B, Finnström O, Lindam A, Nilsson E, Nygren KG, Otterblad Olausson P. Trends in delivery and neonatal outcome after in vitro fertilization in Sweden: data for 25 years. *Human reproduction*. 2010;25(4):1026-1034. doi:10.1093/humrep/deq003

87. Svahn MF, Hargreave M, Nielsen TSS, et al. Mental disorders in childhood and young adulthood among children born to women with fertility problems. *Human reproduction*. 2015;30(9):2129-2137. doi:10.1093/humrep/dev172

88. Rooney KL, Domar AD. The relationship between stress and infertility. *Dialogues in clinical neuroscience*. 2018;20(1):41-47.

89. Alecsandru D, García-Velasco JA. Immunology and human reproduction. *Current opinion in obstetrics and gynecology*. 2015;27(3):231-234. doi:10.1097/GCO.0000000000000174

90. Krieg S, Westphal L. Immune Function and Recurrent Pregnancy Loss. *Seminars in reproductive medicine*. 2015;33(4):305-312. doi:10.1055/s-0035-1554917

91. Grimstad F, Krieg S. Immunogenetic contributions to recurrent pregnancy loss. *Journal of assisted reproduction and genetics*. 2016;33(7):833-847. doi:10.1007/s10815-016-0720-6

92. Mains L, Zimmerman M, Blaine J, et al. Achievement test performance in children conceived by IVF. *Human reproduction*. 2010;25(10):2605-2611. doi:10.1093/humrep/deq218

93. Tavaniotou A, Albano C, Smitz J, Devroey P. Impact of ovarian stimulation on corpus luteum function and embryonic implantation. *Journal of reproductive immunology*. 2002;55(1-2):123-130. doi:10.1016/s0165-0378(01)00134-6

94. Johnson MR, Riddle AF, Grudzinskas JG, Sharma V, Collins WP, Nicolaides KH. Reduced circulating placental protein concentrations during the first trimester are associated with preterm labour and low birth weight. *Human reproduction*. 1993;8(11):1942-1947. doi:10.1093/oxfordjournals.humrep.a137965

95. Haning RV, Goldsmith LT, Seifer DB, et al. Relaxin secretion in in vitro fertilization pregnancies. *American journal of obstetrics and gynecology*. 1996;174(1 Pt 1):233-240. doi:10.1016/s0002-9378(96)70400-3

96. Multiple gestation pregnancy. The ESHRE Capri Workshop Group. *Human reproduction*. 2000;15(8):1856-1864.

97. Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. *International journal of epidemiology*. 2010;39(2):417-420. doi:10.1093/ije/dyp334

98. Knoester M, Helmerhorst FM, Vandenbroucke JP, van der Westerlaken LAJ, Walther FJ, Veen S. Cognitive development of singletons born after intracytoplasmic sperm injection compared with in vitro fertilization and natural conception. *Fertility and*

sterility : official of the American society for reproductive medicine. 2008;90(2):289-296. doi:10.1016/j.fertnstert.2007.06.090

99. Leslie GI, Gibson FL, McMahon C, Cohen J, Saunders DM, Tennant C. Children conceived using ICSI do not have an increased risk of delayed mental development at 5 years of age. *Human reproduction*. 2003;18(10):2067-2072. doi:10.1093/humrep/deg408

100. Maheshwari A, Pandey S, Amalraj Raja E, Shetty A, Hamilton M, Bhattacharya S. Is frozen embryo transfer better for mothers and babies? Can cumulative meta-analysis provide a definitive answer? *Human reproduction update*. 2018;24(1):35-58. doi:10.1093/humupd/dmx031

101. Wennerholm UB, Henningsen AK, Romundstad LB, et al. Perinatal outcomes of children born after frozen-thawed embryo transfer: a Nordic cohort study from the CoNARTaS group. *Hum Reprod*. Sep 2013;28(9):2545-53. doi:10.1093/humrep/det272

102. Roque M, Haahr T, Geber S, Esteves SC, Humaidan P. Fresh versus elective frozen embryo transfer in IVF/ICSI cycles: a systematic review and meta-analysis of reproductive outcomes. *Hum Reprod Update*. 01 01 2019;25(1):2-14. doi:10.1093/humupd/dmy033

103. Fanton M, Nutting V, Solano F, et al. An interpretable machine learning model for predicting the optimal day of trigger during ovarian stimulation. *Fertil Steril*. Apr 30 2022;doi:10.1016/j.fertnstert.2022.04.003

104. Bormann CL, Kanakasabapathy MK, Thirumalaraju P, et al. Performance of a deep learning based neural network in the selection of human blastocysts for implantation. *Elife*. 09 15 2020;9doi:10.7554/eLife.55301

105. Liu L, Jiao Y, Li X, Ouyang Y, Shi D. Machine learning algorithms to predict early pregnancy loss after in vitro fertilization-embryo transfer with fetal heart rate as a strong predictor. *Comput Methods Programs Biomed*. Nov 2020;196:105624. doi:10.1016/j.cmpb.2020.105624

106. Wang CW, Kuo CY, Chen CH, Hsieh YH, Su EC. Predicting clinical pregnancy using clinical features and machine learning algorithms in in vitro fertilization. *PLoS One*. 2022;17(6):e0267554. doi:10.1371/journal.pone.0267554

107. Goyal A, Kuchana M, Ayyagari KPR. Machine learning predicts live-birth occurrence before in-vitro fertilization treatment. *Sci Rep*. 12 01 2020;10(1):20925. doi:10.1038/s41598-020-76928-z

108. Bertini A, Salas R, Chabert S, Sobrevia L, Pardo F. Using Machine Learning to Predict Complications in Pregnancy: A Systematic Review. *Front Bioeng Biotechnol*. 2021;9:780389. doi:10.3389/fbioe.2021.780389

109. Munchel S, Rohrback S, Randise-Hinchliff C, et al. Circulating transcripts in maternal blood reflect a molecular signature of early-onset preeclampsia. *Sci Transl Med*. 07 01 2020;12(550)doi:10.1126/scitranslmed.aaz0131

110. Nair TM. Statistical and artificial neural network-based analysis to understand complexity and heterogeneity in preeclampsia. *Comput Biol Chem*. Aug 2018;75:222-230. doi:10.1016/j.compbiolchem.2018.05.011

111. Axelsson O. The Swedish medical birth register. *Acta obstetricia et gynecologica Scandinavica*. 2003;82(6):491-492. doi:10.1034/j.1600-0412.2003.00172.x

112. Swedish National Board of Health and Welfare. The Swedish Medical Birth Register—A Summary of Content and Quality. Stockholm, Sweden2003. p. 33.

113. Swedish National Board of Health and Welfare. Childbirth after IVF treatment in Sweden 1982–2001. Stockholm, Sweden2006.

114. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish national inpatient register. *BMC public health*. 2011;11:450. doi:10.1186/1471-2458-11-450

115. Wettermark B, Hammar N, Fored CM, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiology and drug safety*. 2007;16(7):726-735. doi:10.1002/pds.1294

116. Statistics Sweden. Multi-generation register 2016. A description of contents and quality. Background Facts, Population and Welfare Statistics (2017:2).

117. Ludvigsson JF, Almqvist C, Ye WM, et al. Registers of the Swedish total population and their use in medical research. *European journal of epidemiology*. 2016;31(2):125-136. doi:10.1007/s10654-016-0117-y

118. Ludvigsson JF, Svedberg P, Olén O, Bruze G, Neovius M. The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research. *Eur J Epidemiol*. Apr 2019;34(4):423-437. doi:10.1007/s10654-019-00511-8

119. Swedish agency for health technology assessment and assessment of social services. ADHD - diagnostics and treatment, organization of the health care and patient involvement. 2013.

120. Jangmo A, Stålhandske A, Chang Z, et al. Attention-Deficit/Hyperactivity Disorder, School Performance, and Effect of Medication. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2019;58(4):423-432. doi:10.1016/j.jaac.2018.11.014

121. Larsson H, Rydén E, Boman M, Långström N, Lichtenstein P, Landén M. Risk of bipolar disorder and schizophrenia in relatives of people with attention-deficit hyperactivity disorder. *Br J Psychiatry*. Aug 2013;203(2):103-6. doi:10.1192/bjp.bp.112.120808

122. Ortqvist AK, Lundholm C, Wettermark B, Ludvigsson JF, Ye W, Almqvist C. Validation of asthma and eczema in population-based Swedish drug and patient registers. *Pharmacoepidemiol Drug Saf*. Aug 2013;22(8):850-60. doi:10.1002/pds.3465

123. Ortqvist AK, Lundholm C, Kieler H, et al. Antibiotics in fetal and early life and subsequent childhood asthma: nationwide population based study with sibling analysis. *BMJ*. Nov 28 2014;349:g6979. doi:10.1136/bmj.g6979

124. Frisell T, Öberg S, Kuja-Halkola R, Sjölander A. Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology*. Sep 2012;23(5):713-20. doi:10.1097/EDE.0b013e31825fa230

125. Sjölander A, Frisell T, Kuja-Halkola R, Öberg S, Zetterqvist J. Carryover Effects in Sibling Comparison Designs. *Epidemiology*. 11 2016;27(6):852-8. doi:10.1097/EDE.0000000000000541

126. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. Jun 29 2009;338:b2393. doi:10.1136/bmj.b2393

127. Raghunathan T, Lepkowski J, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey methodology*. 2001 Jun;27(1):85-96

128. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. Feb 20 2011;30(4):377-99. doi:10.1002/sim.4067

129. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons; 2004 Jun 9.

130. Dreiseitl S, Ohno-Machado L, Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*. 2002;35(5-6):352-359. doi:10.1016/S1532-0464(03)00034-0

131. Suykens JAK, Vandewalle J. Least squares support vector machine classifiers. *Neural processing letters*. 1999;9(3):293-300. doi:10.1023/A:1018628609742

132. Nahar J, Chen YPP, Nahar J, Chen Y-PP, Ali S. Kernel-based naive bayes classifier for breast cancer prediction. *Journal of Biological Systems*. 2007;15(1):17-25. doi:10.1142/S0218339007002076

133. Breiman L. Random forests. *Machine learning*. 2001;45(1):5-32. doi:10.1023/A:1010933404324

134. Friedman JH. Stochastic gradient boosting. *Computational statistics & data analysis*. 2002;38(4):367-378. doi:10.1016/S0167-9473(01)00065-2

135. Vihinen M. How to evaluate performance of prediction methods? Measures and their interpretation in variation effect analysis. *BMC Genomics*. Jun 18 2012;13 Suppl 4:S2. doi:10.1186/1471-2164-13-S4-S2

136. Strobl C, Boulesteix AL, Zeileis A, Hothorn T. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*. Jan 25 2007;8:25. doi:10.1186/1471-2105-8-25

137. Mukund S, Amir N. The Many Shapley Values for Model Explanation. 2020:

138. Strumbelj E, Kononenko I, Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. *Knowledge and information systems : KAIS*. 2014;41(3):647-665. doi:10.1007/s10115-013-0679-x

139. Lipovetsky S, Conklin M, Lipovetsky S, Conklin M. Analysis of regression in game theory approach. *Applied stochastic models in business and industry*. 2001;17(4):319-330. doi:10.1002/asmb.446

140. Wagenaar K, van Weissenbruch MM, van Leeuwen FE, et al. Self-reported behavioral and socioemotional functioning of 11- to 18-year-old adolescents conceived by in vitro fertilization. *Fertil Steril*. Feb 2011;95(2):611-6. doi:10.1016/j.fertnstert.2010.04.076

141. Rissanen E, Gissler M, Lehti V, Tiitinen A. The risk of psychiatric disorders among Finnish ART and spontaneously conceived children: Finnish population-based register study. *Eur Child Adolesc Psychiatry*. Aug 2020;29(8):1155-1164. doi:10.1007/s00787-019-01433-2

142. Wijs LA, Fusco MR, Doherty DA, Keelan JA, Hart RJ. Asthma and allergies in offspring conceived by ART: a systematic review and meta-analysis. *Hum Reprod Update*. Dec 21 2021;28(1):132-148. doi:10.1093/humupd/dmab031

143. Rodriguez-Wallberg KA, Lundberg FE, Ekberg S, et al. Mortality from infancy to adolescence in singleton children conceived from assisted reproductive techniques versus naturally conceived singletons in Sweden. *Fertil Steril*. Mar 2020;113(3):524-532. doi:10.1016/j.fertnstert.2019.10.018

144. Rodriguez-Wallberg KA, Berger AS, Fagerberg A, et al. Increased incidence of obstetric and perinatal complications in pregnancies achieved using donor oocytes and single embryo transfer in young and healthy women. A prospective hospital-based matched cohort study. *Gynecol Endocrinol*. Apr 2019;35(4):314-319. doi:10.1080/09513590.2018.1528577

145. Pinborg A, Henningsen AA, Loft A, Malchau SS, Forman J, Andersen AN. Large baby syndrome in singletons born after frozen embryo transfer (FET): is it due to maternal factors or the cryotechnique? *Hum Reprod*. Mar 2014;29(3):618-27. doi:10.1093/humrep/det440

146. Hiura H, Hattori H, Arima T, et al. Genome-wide microRNA expression profiling in placentae from frozen-thawed blastocyst transfer. *Clinical epigenetics*. 2017;9(1)doi:10.1186/s13148-017-0379-6

147. Reyes Palomares A, Rodriguez-Wallberg KA. Update on the Epigenomic Implication of Embryo Cryopreservation Methods Applied in Assisted Reproductive Technologies With Potential Long-Term Health Effects. *Front Cell Dev Biol*. 2022;10:881550. doi:10.3389/fcell.2022.881550

148. Roque M, Valle M, Guimarães F, Sampaio M, Geber S. Freeze-all policy: fresh vs. frozen-thawed embryo transfer. *Fertil Steril*. May 2015;103(5):1190-3. doi:10.1016/j.fertnstert.2015.01.045

149. Blockeel C, Drakopoulos P, Santos-Ribeiro S, Polyzos NP, Tournaye H. A fresh look at the freeze-all protocol: a SWOT analysis. *Hum Reprod*. Mar 2016;31(3):491-7. doi:10.1093/humrep/dev339

150. Bell JL, Hardy P, Greenland M, et al. E-Freeze - a randomised controlled trial evaluating the clinical and cost effectiveness of a policy of freezing embryos followed by thawed frozen embryo transfer compared with a policy of fresh embryo transfer, in women undergoing in vitro fertilisation: a statistical analysis plan. *Trials*. Jun 30 2020;21(1):596. doi:10.1186/s13063-020-04441-9

151. Wong KM, van Wely M, Verhoeve HR, et al. Transfer of fresh or frozen embryos: a randomised controlled trial. *Human reproduction*. 2021;36(4):998-1006. doi:10.1093/humrep/deaa305

152. Roberts JM, Cooper DW. Pathogenesis and genetics of pre-eclampsia. *The Lancet*. 2001;357(9249):53-56. doi:10.1016/S0140-6736(00)03577-7

153. De Kat AC, Hirst J, Peters SA, et al. Prediction models for preeclampsia: A systematic review. *Pregnancy hypertension*. 2019;16:48-66. doi:10.1016/j.preghy.2019.03.005

154. Neary C, Naheed S, McLernon DJ, Black M. Predicting risk of postpartum haemorrhage: a systematic review. *BJOG : an international journal of obstetrics and gynaecology*. 2021;128(1):46-53. doi:10.1111/1471-0528.16379

155. Bernsen RM, de Jongste JC, van der Wouden JC. Birth order and sibship size as independent risk factors for asthma, allergy, and eczema. *Pediatr Allergy Immunol*. Dec 2003;14(6):464-9. doi:10.1046/j.0905-6157.2003.00108.x

156. Cawley GC, Talbot NLC. On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. *Journal of machine learning research : JMLR*. 2010;11:2079-2107.