Discussion
The aim of this thesis is to analyse the possible effect of different IVF components, i.e., COH, the in vitro culture procedure, and PGS, and the underlying subfertility on long-term child health and development. Ovarian hyperstimulation induces the growth of multiple follicles, bypassing the natural selection of the development of one dominant follicle, and leads to higher progesterone and oestrogen levels preceding the laboratory IVF procedures. During the in vitro culture procedure and PGS oocytes, sperm and embryos are cultured in vitro, possibly affecting their developmental phenotype by epigenetic reprogramming. In addition, in PGS the embryo is manipulated even further than in IVF. As it is known that the early environment of an embryo may affect its physiology and thereby may lead to an increased vulnerability for health related problems in later life, it is conceivable that ovarian stimulation, the in vitro culture procedure and PGS affect the health of the offspring. In addition, the aetiology of subfertility or the parental subfertility may contribute to the risk of adverse health outcomes in the offspring, even though the etiological pathways are not clear.

Knowledge on the possible associations between parts of the IVF procedures, PGS, subfertility and child health is important since IVF and PGS are increasingly used worldwide. In the discussion first the main findings in the Groningen ART cohort in relation to the current literature are discussed. Next, the discussion continues by addressing the findings in the PGS-trial follow-up study, followed by methodological limitations. Future perspective and concluding remarks close the discussion.

Discussion of the results - Groningen ART Cohort study

Reflection on the results

Twins and singletons born after IVF

Our study demonstrated that 4-year-old IVF conceived twins had a lower height (2.9 cm), a lower weight (1.7 kg) and a lower IQ (5.4 points) as compared to IVF singletons and a similar neurodevelopment and BP. Yet, after adjustment for confounders and mediators the development and health of 4-year-old IVF conceived twins did not differ from that of 4-year-old IVF conceived singletons. This indicates that the less favourable IQ scores, height and weight of IVF conceived twins is largely mediated by the increased risk of twins for adverse perinatal outcome and is not altered by IVF per se. An increased risk of perinatal adversities was observed in the IVF twin group compared to the IVF singleton group. This finding is in line with that of others who studied the perinatal outcomes of IVF conceived twins and singletons. In addition, it is also in line with the large body of literature on naturally conceived twins and singletons. Although an increased risk of perinatal adversities was observed in IVF twin pregnancies compared to IVF singleton pregnancies, only a few studies addressed the long-term developmental outcomes after the perinatal period in twins born after IVF. Our findings regarding neurodevelopment, cognitive development and anthropometrics are in line with these studies. To the best of our knowledge no other study investigated the BP levels of twins and singletons born after IVF. This lack of knowledge about health and development of IVF twins emphasizes the need of good quality studies which track the health of ART offspring in long-term follow-up.

Still, our study could have been improved by adding a group of naturally conceived twins. Comparing IVF conceived singletons and IVF conceived twins with naturally conceived twins could tell us if IVF increases the risk for a less optimal health and development in twins.

Conclusion:

Our study indicated that the generally observed less favourable health and development in twins is also present in IVF conceived twins compared to IVF conceived singletons. However, our findings are not conclusive and should be interpreted with caution, as we do not know whether IVF actually increases this risk, as we did not compare naturally conceived twins with IVF conceived twins.
Clinical implication:
In line with other studies our study indicated that a singleton pregnancy should be favoured over a twin pregnancy regarding short and long-term outcomes and that therefore a policy of a single embryo transfer is recommended when appropriate. In the Netherlands a decrease in the twin delivery rate after IVF and ICSI is observed from 2003 (803/3706 [21.7%]) up until 2017 (155/5067 [3.1%]). In other Nordic-European the twin delivery rate after IVF and ICSI decreased as well. Yet, the twin delivery rate in most other European regions increased.

IVF and the risk of asthma and atopic disease in the offspring
At the age of 4 years COH-IVF offspring more often used asthma medication than Sub-NC offspring. This finding differs from our finding at the age of 9 years. At the age of 9 years the IVF procedures (COH and the in vitro culture procedure) were not associated with an increased risk of asthma in the offspring. This difference in findings possibly could be attributed to the difference in age (4 versus 9 years) and be due to the fact that we used two different questionnaires.

At pre-school age (3-to-4 year olds) it is relatively hard to diagnose asthma, as spirometry cannot be used, whereas at the age of 9 years spirometry can be used. When children had reached the age of 4 years the parents were asked two questions: a) does your child have the diagnosis asthma, and b) does your child use asthma medication? In the Netherlands specific guidelines are used to diagnose asthma under the age of six. Because all the doctors use the same standard to diagnose asthma, the diagnosis of asthma should be accurate. However, it would have been better if the medical professionals in charge of the children with pulmonary signs had collected standardized data about the children's pulmonary condition, which we (after permission of the parents) collected. In addition, medication may also have been prescribed for other indications than asthma, such as viral infections. Note, that our questionnaire did not include a question on the severity of asthma.

At the age of 9 years we used the standardized and validated International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire. Because we used a standardized and validated questionnaire at the age of 9 years we consider the findings at 9 years more valid. In addition asthma symptoms that start in early childhood may disappear and may be absent for a relatively long time in life. Note that the predisposition for asthma itself does not resolve. However, the diminished expression of asthma with increasing age could also explain the difference in our findings. Still, the asthma prevalences in the three groups were lower than in the general population and lower than in other studies addressing the prevalence of asthma in IVF offspring. This underlines the notion that our study groups are not representative of the general population.

Other studies evaluating the potential association between ART and asthma in the offspring suggested an increased risk for asthma in ART offspring, but could not determine if this was due to one of the IVF procedures, subfertility per se or confounding variables. Within the Millennium cohort study (n= 18,818) Carson and colleagues found an increased prevalence of asthma in 5-year-old children conceived with ART compared to naturally conceived children to fertile couples: adjusted OR (95% CI): 2.38 (1.34, 4.24) (after adjustment for breastfeeding, Caesarean section, gestational, maternal history of asthma, family smoking, social class and maternal age). At the age of 7 years still an increased asthma prevalence was observed: (adjusted OR [95% CI]: 1.84 [1.03, 3.28]), but no adjustment was made for breastfeeding, Caesarean section, gestational age or low birthweight. This leaves us with the question what causes the increased risk for asthma in ART offspring: our study indicated that the increased risk for asthma in IVF offspring is probably not due to the IVF procedures.

Within the Swedish birth register cohort (n=2,628,728) Källén et al. used asthma drug use as an outcome parameter to study the prevalence of asthma in IVF offspring aged 2 to 25 years. They demonstrated an increased risk for asthma in children conceived with IVF: adjusted OR (95% CI) 1.28 (1.23, 1.34). Yet, and in line with our study, this increased risk could not be attributed to the IVF procedures. The authors concluded that the risk could be attributed to parental subfertility, with neonatal morbidity and maternal asthma acting as mediators. The study did not adjust for low birthweight or Caesarean section, which are known risk factors for the development of asthma.

The lack of information on confounding factors is a drawback of multiple studies that investigated the associations between IVF, parental subfertility and asthma in the offspring. Besides the known risk factor, there are several suspected but unverified risk factors for the development of asthma. For example, maternal high dose folic acid use (5 mg) during pregnancy might increase the risk of asthma in the offspring. In general, studies do not have information on these suspected but unverified risk factors.

Conclusion:
No direct link between the IVF procedures per se and asthma in the offspring has been established in our study of 9-year-old children born to subfertile couples. Some studies showed an effect of parental subfertility on the prevalence of asthma in IVF offspring. However, it is unclear how this mechanism (parental subfertility leading to asthma in the offspring) works. It is possible that confounding factors cause the association.

Clinical implication:
Asthma is a disorder with a considerable impact, not only in terms of reduced health but also in terms of costs. For instance, in Europe the estimated mean costs per patient per
Cardiovascular health of IVF offspring

The BP levels of 9-year-old singletons born following COH-IVF, MNC-IVF and Sub-NC did not differ in our study. Our finding differs from a previous finding of our study group. At the age of 4 years IVF children born following COH-IVF had a higher SBP than children born following MNC-IVF, suggesting an unfavourable effect of ovarian hyperstimulation on the offspring’s BP. As the literature describes a linear relationship between pre-school age BP and school-age BP the disappearance of the difference in BP between the COH-IVF and MNC-IVF group is a remarkable finding. A possible explanation could be that COH-IVF offspring is more vulnerable to stress than MNC-IVF offspring. It is known that the chances of achieving a pregnancy after one MNC-IVF cycle are smaller than after one COH-IVF cycle. Therefore women often have to undergo multiple MNC-IVF cycles. It is possible that women who chose to undergo a MNC-IVF trajectory tend to be more emotional stable (cope well with stress) and more perseverant, as this is a longer trajectory compared to a COH-IVF trajectory. These two character traits potentially could have a positive influence on the offspring by nature and nurture. It is conceivable that the BP-assessment at 4 years had functioned as a rather stressful condition, as at the age of 4 years the child’s cognitive abilities to understand the strange situation of a BP-measurement are limited. At 9 years, children are presumably less stressed by a BP-measurement, as their cognitive abilities allow them to classify the measurement as a somewhat weird but harmless medical test.

Furthermore, at the age of 9 years just like at the age of 4 years, the mean BP percentiles of the three groups of children born to subfertile couples were higher than the expected 50th percentile. This indicates that the 9-year-old offspring of subfertile couples could have a higher BP than children who are conceived naturally. A second study of our group showed that the BP percentiles of 9-year-old offspring born at term of subfertile couples have a higher BP than children born following MNC-IVF, suggesting an unfavourable effect of ovarian hyperstimulation on the offspring’s BP. This finding is in line with others reporting that the offspring of subfertile couples has a higher BP than adults born to fertile couples. In case the offspring of subfertile couples have a higher BP than adults born to fertile couples. In case the offspring of subfertile couples have a higher BP at adult age cardiovascular risk management should take being born to subfertile couples into account.

Visual acuity in children conceived with IVF

Our study showed that the proportion children wearing glasses on medical indication was highest in the COH-IVF group (23%), followed by the MNC-IVF (12%) and the Sub-NC (8%) groups. After adjustment the proportion of COH-IVF children wearing glasses on medical indication was four times higher than that of Sub-NC children (adjusted odds ratio 4.0 [95%CI: 1.13, 14.06]) and two times higher than that of MNC-IVF children (adjusted odds ratio 2.4 [95%CI: 0.74, 7.87]). This indicates that the observed differences may be related to the COH used in IVF.

The Dutch prevalence of 11-year-old children wearing glasses, which is 11%, supports our finding, as the prevalence of wearing glasses on medical indication in the COH-IVF reasons (it will cost more money and time) not possible.

Other studies showed that children conceived by IVF/ICSI had statistically significantly higher BP levels than naturally conceived children. However, the studies compared IVF/ICSI offspring with naturally conceived children to fertile couples. Without adjustment for subfertility, the higher BP levels could also be attributed to subfertility.

Others who did have information on subfertility demonstrated higher BP levels in 5-to-6 year old children born to subfertile couples compared to children born to fertile couples. This finding is in line with others reporting that the offspring of subfertile couples have a less optimal health compared to naturally conceived children of fertile couples, such as adverse perinatal outcomes and birth defects.

Conclusion:

Controlled ovarian hyperstimulation and the in vitro procedure are not associated with higher BP in 9-year-old children born to subfertile couples. Nine-year-old children of subfertile couples have higher BP levels than naturally conceived children of fertile couples. The higher BP levels may be attributed to the multifactorial inherent risk associated with subfertility and not to the IVF procedures.

Clinical implications:

High BP is one of the main risk factors for the development of a cardiovascular disease and can damage the human body years before the first symptoms of a cardiovascular disease emerge. As it is known that BP tracks from childhood into adulthood, these differences in BP between fertile and subfertile offspring may become more evident when the children grow older. Therefore it is important to investigate if adults born to subfertile couples have a higher BP than adults born to fertile couples. In case the offspring of subfertile couples have a higher BP at adult age cardiovascular risk management should take being born to subfertile couples into account.

Discussion |
The prevalence of children in the MNC-IVF group and Sub-NC group wearing glasses was comparable with the Dutch average rate of children wearing glasses. This could mean that COH per se is associated with refractive errors in IVF offspring.

However, our study only had information on children’s ophthalmic condition that was obtained from their ophthalmologist. It is preferred that all participating children undergo the same ocular examination that includes: visual acuity, refraction, cornea thickness, axial length, ocular surface and intraocular pressure. By doing so, it is possible to investigate if COH is also associated with other ocular outcomes.

Up until now one systematic review, conducted in November 2018, evaluating the visual acuity of children conceived with the help of ART has been carried out. The review identified six studies examining the visual function of ART offspring. Their main finding was that the available data were inadequate to draw a firm conclusion regarding IVF and the visual function of the offspring: sample sizes in all but one study were insufficient and multiple studies lacked information about the obstetric and perinatal period. The one large observational based Swedish study of Tornqvist et al. demonstrated an increased incidence of severe visual impairment in 24,628 IVF children born between 1985-2005 in Sweden, with all children born during the corresponding years being used as the control group (odds ratio [95% CI]: 1.65 [1.12, 2.45]). This finding is in line with our results.

**Conclusion:**
Our study found support for the hypothesis that IVF offspring is at risk for refractive errors at the age of 11 years. This observed risk might be related to the COH used in IVF.

**Clinical implications:**
No study has been able to confirm or reject the hypothesis that IVF offspring is vulnerable for an altered ophthalmic development. In the Netherlands, the general health, including the ophthalmic condition, of all children up until pre-school age is evaluated multiple times at the well-baby clinics (Consultatie Bureau). If our result can be replicated in well-powered studies special attention should be given to the visual acuity of IVF offspring at school age.

---

**Discussion of the results – the PGS-trial**

**Developmental outcome of children born after PGS**

In the follow-up study of our multicentre RCT PGS in cleavage stage embryos was not associated with adverse effects on neurological, cognitive and behavioural development, BP or anthropometrics in IVF offspring at 9 years. This finding differs from two previous findings of our research group, in which we reported that (a) 2-year-old PGS offspring had a lower neurological optimality score than controls and (b) 4-year-old PGS twins had a lower neurological optimality score than control twins. This difference in findings might be attributed to plasticity of the brain in childhood that result in the disappearance of differences in the neurological optimality score. Alternatively, it might be that our group sizes were not sufficiently large to detect minimal effects of PGS on the sensitive neurological optimality score.

A remarkable finding of our study was the relatively high prevalence of an adverse neurological outcome, i.e., the presence of minor neurological dysfunction (MND), at the age of 9 years. In our study the prevalence of an adverse neurological outcome was 36% (PGS 40% and control group 34%), whereas the prevalence of an adverse neurological outcome (MND) in the general population is 5-7%. Olivares et al. recently concluded that parental subfertility is associated with a higher prevalence of an adverse neurological outcome in the offspring, not the specific components of IVF.

Our findings on cognitive and neuromotor development are in line with others. The observational study of Winter et al. did not find a difference in psychomotor development between 5-to-6-year-old day-3 PGD offspring and ICSI offspring. Our findings on BP and anthropometric outcomes do not differ from our findings at the age of 4 years and are in agreement with the study of Belva et al. on the body composition and BP of 5.5-year-old day-3 PGD offspring. Belva and colleagues also reported no differences in standing height, weight, body fat distribution or BP between PGD and ICSI offspring. However, it should be noted that in our study the mean BP percentiles of the PGS and control group were above the 60th percentile. In line with other reports (see section on cardiovascular health) this suggests that IVF offspring born with and without PGS are vulnerable to an increased BP.

However, as our previous study on IVF and BP we did not have information on physical activity and the diet of the participation children and all BP measurements were performed on one day. Other limitations were discussed in the chapter itself (selective drop-out, power calculation and day of embryo biopsy).

**Conclusion:**
Day-3 cleavage-stage PGS does not have adverse consequences for neurodevelopment and cardiovascular health of 9-year-old IVF offspring.
Clinical implications:
That day-3 cleavage-stage embryo biopsy does not have adverse consequences for neurodevelopment and cardiovascular health of the offspring is essential information for couples with a high risk of transmitting a genetic disorder that request PGD. Although day-3 embryo biopsy is no longer recommended, day-3 cleavage stage embryo biopsy for PGD and PGS continues to be performed. It is reassuring for both patients and physicians that day-3 embryo biopsy does not affect child health up until the age of 9 years.

Methodological considerations
We acknowledge limitations within our study. One of the main limitations of the study is its size. Initially the study focused on the neurological development of children born after IVF with and without COH and children who were conceived naturally to subfertile couples. The children’s neurological development was evaluated with the Hempel assessment. As mentioned in Box 1 (page 30) one of the outcome measures of the Hempel examination is the fluency score, which is sensitive and suitable to find a dysfunction in the nervous system of children. Based on the fluency score at 18 months 64 children had to be included in each group to reach a power of 80%. The limited sample sizes did not allow us to stratify our data by ICSI and sex of the children. Although the major strength of the study is its unique design, the design does not allow us to study the possible effect of subfertility per se. The inclusion of a group of children born to fertile couples at birth would have improved the study’s design. Note, that a proper design would have included the children at birth; retrospective inclusion of fertile controls induces selection bias. Our study had detailed information on TTP that can be used as a proxy for the severity of subfertility. However, there could be a difference in the meaning of TTP in the IVF groups and the Sub-NC groups. In the former, TTP ends by the performance of IVF and one never will know if these couples would get spontaneously pregnant if no IVF was used. So a real TTP is actually unknown in these groups, as IVF has shortened TTP artificially.

Another limitation is that the women who underwent MNC-IVF treatment had to fulfill specific inclusion criteria regarding age, BMI, previous type of ART treatment and menstrual cycle. Although we have tried to reduce selection bias by statistical adjustment we cannot rule out this form of bias. Other subter group characteristics could possibly also influence health and development of the offspring. For example, the chances of achieving a pregnancy after one MNC-IVF cycle are smaller than after one COH-IVF cycle. Therefore women often have to undergo multiple MNC-IVF cycles. It is possible that women who chose to undergo a MNC-IVF trajectory tend to be more emotional stable (cope well with stress) and more perseverant, as this is a longer trajectory compared to a COH-IVF trajectory. These two character traits potentially could have a positive influence on the offspring by nature and nurture.

Future perspectives
Scientists have been asking themselves the legitimate question if we are not overusing IVF and whether there are long-term consequences of IVF for the offspring. Both questions do not detract from the fact that over seven million children have been conceived with the help of IVF and that little is known about the long-term health outcomes of ART offspring. To gain knowledge about IVF offspring’s health, follow-up studies and population-based studies evaluating the long-term safety of IVF are needed. However, long-term follow-up studies in general, as illustrated by this thesis, are difficult to carry out and population-based studies require financial plus political support. However, with ART birth records and big data analysis becoming more common hopefully these types of studies can be carried out easier in the near future. Another obstacle that the scientific community addressing the long-term sequelae of ART has to tackle is that the ART-procedures develop and change over time. Therefore ART birth records should register the specifics of the ART procedure including its indication and the details of the procedure, such as type of hyperstimulation, types of medication, duration of culture, ICSI or IVF, type of culture media, air-condition, so that the data can be linked to data on health of ART offspring in the future.

Multiple issues that need to be addressed in future studies were already mentioned throughout this thesis. Nevertheless I want to highlight three research topics.
(1) Futures studies should address the question what the influence of the IVF procedures and subfertility per se are on the cardiovascular and cardiometabolic health of the offspring. Within the Groningen ART cohort and PGS trial follow-up study we aim to conduct another follow-up round at the age of 18 years. Lipid profiles, glucose and insulin levels could be added to the measurement of BP, anthropometrics and advanced glycated end products. A benefit of the measurement of these metabolic parameters is the availability of reference values allowing for a comparison with the values of the general population.
(2) Evidence is increasing that parental subfertility has a negative influence on child health. Studies on the long-term health of ART offspring showed that after adjustment for TTP most adverse outcomes are reduced. However, there is a downfall of taking TTP as a proxy for the severity of subfertility. In subfertile and fertile control groups women conceived naturally, i.e. their TTP naturally ended, whereas in ART groups pregnancy was achieved after a medical intervention, i.e. their TTP is artificially ended. This means that there always could be a remaining effect of subfertility on ART offspring after adjustment. Studies consisting of a group of naturally conceived children born to subfertile couples who were waiting for fertility evaluation or treatment and compare this group of children to naturally conceived children to fertile couples could give more insight on the effect of parental subfertility on child health. A RCT on IVF with a no
intervention arm’ could be of great value. But this introduces great ethical issues. It would be interesting to see if in the ‘no intervention arm’ the TTP - which is a proxy for the severity of subfertility - has an effect on the health and development of the offspring. Due to the common use of IVF and the change of the definition of subfertility from 2 years to 1 year by the World Health Organisation in 2008 there is a tendency to treat less severe subfertile couples. For example it is possible that the more favourable outcomes regarding preterm birth described in the introduction could be explained by the inclusion in the treatment of less severe subfertile couples and not by improved IVF techniques.

In addition, most studies investigate child health within a mixed population of maternal subfertility, paternal subfertility and unexplained subfertility. Future studies should address if maternal and paternal subfertility have independent or combined effects on child health.

(3) Another issue that needs to be addressed is the reproductive health of ART offspring.

Louise Brown’s younger sister, Natalie, was the first female born after IVF to give birth herself. Regarding the reproductive health of their IVF-peers little is known.

Most studies assessed the reproductive abilities of male ICSI offspring (ICSI used as treatment for male infertility diagnosis) and found that in 54 ICSI men aged 18-to-22-years more often sperm concentrations below 15 million/mL and a total sperm counts below 39 million than naturally conceived peers.38 Men with sperm concentrations below 15 million/mL and/or a total sperm counts below 39 million are considered to have a low sperm count.39 However, little is known on the effect of TTP on outcome of ICSI offspring as no study addressed this issue. Still these first results in a small group of adult men born after ICSI emphasize the need of studies on the reproductive abilities of ART offspring.

Concluding remarks

Overall, the results of our studies regarding the IVF procedures and PGS are reassuring. We found no differences in cardiovascular health and asthma prevalence between singletons conceived after IVF with and without ovarian hyperstimulation and naturally conceived children to subfertile couples at the age of 9 years. In addition, day-3 PGS does not seem to have an adverse effect on child health. However we did find that children conceived with conventional IVF, i.e. IVF with ovarian hyperstimulation, are at risk for refractive errors and that subfertility is associated with a less favourable cardiovascular risk profile. In the current literature evidence is accumulating that the majority of adverse effects found in ART offspring are due to the underlying parental subfertility and not to the ART procedures. This means that ART allows subfertile couples to conceive of offspring that may be at risk of an impaired health at later age. Further studies should determine how this risk is at the individual and population level. Therefore, follow-up studies addressing the long-term health outcomes after ART are needed.

The five main findings of this thesis are:

(I). The generally observed increased risk for less favourable health in twins also holds true for IVF twins compared to IVF singletons.

(II). The IVF procedures are not associated with an increased prevalence of asthma and rhinitis in the offspring at 9 years of age.

(III). Ovarian hyperstimulation as part of an IVF-procedure is associated with refractive errors in 11-year-old IVF offspring.

(IV). The IVF-procedures are not associated with higher blood pressure values in 9-year-old children born to subfertile couples, subfertility per se is.

(V). Preimplantation genetic screening following day-3 cleavage stage embryo biopsy does not have an adverse effect on the neurodevelopment, anthropometrics and blood pressure of 9-year-old IVF offspring.
References


