CHAPTER 8
General discussion
This thesis evaluates the effects related to assisted conception on neurodevelopmental outcome in preschool-aged children. Standardized, age-specific and sensitive tools were applied to study neurodevelopmental outcome in the participants of two parallel running projects, the Groningen ART (assisted reproductive techniques) cohort study and the PGS (preimplantation genetic screening) study. In general, this thesis demonstrated that ART-related aspects such as ovarian hyperstimulation, the in vitro laboratory procedures, or a combination of both were not associated with worse neurodevelopmental outcome – in terms of movement variation, neuromotor function, cognition and behaviour – in singleton children aged 4 months to 4 years. Additionally, a history of subfertility per se and the underlying cause of subfertility were not associated with worse neurodevelopmental outcome at age 2 and 4. However, increased time to pregnancy (TTP) was associated with an increased risk for minor neurological dysfunction (MND) at 2 years and an increased risk for the complex form of MND at 4 years. Moreover, direct negative effects were found between both the severity of subfertility in terms of TTP and cognition and the presence of subfertility and behaviour at age 4.

In this thesis it was also observed that neurodevelopmental outcome – in terms of neuromotor function, cognition and behaviour – in singletons born following in vitro fertilization (IVF) with PGS is similar to that of singletons born after IVF without PGS at age 4. However, PGS may be associated with altered neurodevelopment in twins.

This chapter starts with a discussion of the main findings of the studies conducted. First, the main findings of the Groningen ART cohort study will be discussed, in terms of ART-related and subfertility-related aspects related to child neurodevelopmental outcome. Hereafter, the chapter focusses on the implications of ART and subfertility for daily practice and the consequences of minor neurological signs. Next, the main findings of the PGS study, will be discussed, including the clinical implications of embryo biopsy. The chapter continues with an overview of methodological considerations per project, some suggestions for future research and ends with the concluding remarks.

**ART aspects related to child neurodevelopmental outcome**

This thesis concludes that rather the severity of subfertility than its presence or underlying cause or ART-related aspects such as ovarian hyperstimulation and the in vitro procedures affect neurodevelopmental outcome in preschool-aged singleton children, born to subfertile couples. This finding corroborates previous findings within the Groningen ART cohort study: similar neurodevelopmental outcome between the three ART study groups were found at 2 weeks and 3, 4, 10 and 18 months of age. Moreover, our observation that ART-related aspects do not negatively affect the child’s neurodevelopmental outcome correspond to the majority of follow-up studies. Does this mean that ART is entirely safe for the developing child? Although generally no significant differences were found between COH-IVF, MNC-IVF and Sub-NC children, a closer look at the results in this thesis learns that once a difference was found between group COH-IVF and MNC-IVF. In the period
of 4 to 18 months of age, COH-IVF children had lower IMP variation scores than MNC-IVF children. An explanation for the finding that ART affects the IMP variation score rather than the prevalence of MND or the optimality scores may be the fact that variation in motor behaviour taps on other neural circuitries than traditional neurological signs. Variation in motor behaviour in particular reflects integrity of cortical connectivity. A similar specific effect of ovarian hyperstimulation was observed in another study that was part of the current Groningen ART cohort study in terms of cardiometabolic outcome. Seggers et al. and La Bastide-Van Gemert et al., demonstrated a direct negative effect of ovarian hyperstimulation on systolic blood pressure percentiles and subscapular skinfold thickness at age 4. A potentially adverse effect of ovarian hyperstimulation could be mediated by an altered oocyte quality, due to the production of multiple follicles by the ovaries, resulting in non-optimal oxygenation of the rapidly growing follicles and the bypass of the dominant follicle. Furthermore, both animal and clinical and studies suggest that supraphysiological estradiol levels may be toxic for the developing embryo by impairing its implantation potential. Ovarian hyperstimulation may also interfere with the intra-uterine environment. The alterations in hormonal levels may lead to disturbances in the endometrium, resulting in diminished endometrial receptivity, a non-optimal expression of endometrial growth and thus, a poorer implantation environment. The altered early environment could induce a non-optimal development of the oocyte and embryo. A study by Pelinck et al. supports the latter thought. The authors reported that COH-IVF children had a grams lower birthweight compared to MNC-IVF, suggesting that ovarian hyperstimulation may be a causative factor rather than the in vitro procedure per se in the occurrence of low birthweight after conventional IVF.

However, an argument against an unfavourable effect of ovarian hyperstimulation is the finding that the variation scores of children born following COH-IVF did not differ significantly from that of Sub-NC children. Moreover, for all of the other neurodevelopmental outcome parameters, no differences were found between COH-IVF children and Sub-NC children during the whole follow-up period from 4 months of age until 4 years of age. Many other studies also conclude that early neurodevelopmental outcome of children born after IVF – i.e. the commonly applied IVF with ovarian hyperstimulation – does not differ from children conceived naturally (NC).

Subfertility aspects related to child neurodevelopmental outcome

This thesis concludes that the presence of a history of subfertility per se or the underlying cause of subfertility do not seem to affect the neuromotor development of children at the age of 2 and 4 years. At 2 years, no differences in optimality scores or the prevalence of MND could be observed between the children born to the Groningen ART cohort and children born to a fertile reference group. Similar results were found for the optimality scores and prevalence of complex MND at 4 years. These are initially reassuring findings, however, the absence of a difference in outcome between children born to subfertile parents and those born to fertile parents may be explained by the composition of the
reference group. An argument for this hypothesis lies in the study of Middelburg et al. on neurodevelopmental outcome of the Groningen ART cohort children at 3 months. In this study, a significantly worse neurological outcome was found for the cohort children compared to a reference group that consisted of a representative sample of the general population, as the infants had been assessed as a part of a general health check-up provided for all infants. At the follow-up at 2 and 4 years – when no significant neurological disadvantage was found for the subfertile group – the fertile reference group consisted of children whose parents volunteered for participation in the study. This may have caused selection bias. Perhaps, parents who are more concerned about the health and development of their child are more willing to participate in a neurodevelopmental outcome study. The idea of selection bias is supported by the relatively high proportion of children in the fertile reference group with complex MND (14%), a proportion that is higher than that reported in the general population (6-7%). A similar phenomenon was seen in the study by Knoester et al. That study reported a higher prevalence of MND in the control group of NC children compared to the general population. Given this potential selection bias, the slightly worse neurodevelopmental outcome the authors found in their 5 to 8-year-old ICSI children compared to their NC children may have actually been larger.

The composition of the fertile reference group may be an explanation for the absence of an effect of subfertility per se on child neuromotor outcome at 2 and 4 years. However, in the study on the ART treatment effects and subfertility effects on cognition and behaviour in 4-year-olds (Chapter 6), a direct negative causal effect was found between the presence of subfertility and behaviour, an effect that came to light with the inclusion of the fertile reference group. Also, a direct negative causal effect was found of the severity of subfertility, reflected by a longer TTP, and the child’s cognition. Both effects were confounded by maternal age at child conception and maternal educational level and cognition and behaviour were directly related to one another. This implies that suffering from subfertility per se, and especially from more severe subfertility – which by itself is associated with higher age and high educational level of the mother – negatively affects the child’s cognitive and behavioural outcome. A similar finding was demonstrated in Chapter 4 and 5, where it is concluded, that a longer TTP was associated with a less optimal neurological condition in preschool-aged children. The effect of TTP on neurodevelopmental outcome was investigated at age 2 and age 4. At 2 years of age, TTP was associated with a higher prevalence of MND, whereas at 4 years of age it was associated with a higher prevalence of complex MND. In the 4-year follow-up round, special attention was paid to potential sex-differences in outcome. The negative association found in the study was particularly true for girls. In general, boys have twice as often complex MND as girls, which may be related to the sex-specific differences and timing of the development and maturation of the nervous system and, with that, the sex-depended critical period for the occurrence of developmental events. However, in our subfertile study group a relatively high proportion of children, boys as well as girls, had complex MND (total group: 24%; boys: 21%; girls: 28%), a proportion that is considerably higher than in the
general population (6-7%). One could wonder why a longer TTP was only associated with clinical neurological condition in terms of complex MND and not with the derived parameters fluency score and NOS. Clinical neurological condition and neurological optimality are highly correlated; however, a reduction in neurological optimality does not necessarily mean the presence of a neurological deviation, as the range for optimality is narrower than that for normality. Moreover, the classification of MND involves a threshold: MND is only present if a certain number of deviant neurological signs are present. This means that the severity of subfertility is not associated with milder degrees of neurological non-optimality, but only has an impact on the risk for the clinically relevant form of MND.

In line with the result of the current thesis, some studies demonstrated that subfertility-related aspects rather than ART treatment aspect may contribute to a worse child neurodevelopmental outcome. Zhu et al. found a delay in achieving certain motor milestones in 18 month year-old ICSI children and a modest increased risk for developmental coordination disorder (DCD) in 7-year-old children born to subfertile couples. However, the scarce literature available is less supportive towards the finding that the severity of subfertility may be the decisive factor rather than the presence of a history of subfertility per se. For instance, Zhu et al. were not able to demonstrate an effect of prolonged subfertility on the prevalence of behavioural problems in 7 to 21-year-old offspring. Carson et al. could not demonstrate an adverse effect of subfertility on children's cognitive development at age 3 or 5, nor did they find evidence that a prolonged period of subfertility was associated with adverse child behaviour at age 5 and 7. Children born after ART had a higher mean total difficulties score than their NC peers, yet it seems that their scores mostly remained within the normal range. The authors speculate that the reported behavioural problems by the ART parents may be an underestimation of real problems as generally, ART parents tend to cast their children in an overly positive light and report greater warmth and parental involvement.

A history subfertility, severe subfertility and child neurodevelopmental outcome
From the results of the studies in this thesis, the question arises why subfertility-related aspects, especially the severity of subfertility, may affect the child's neurodevelopmental outcome. Several explanations for the finding will pass in review.

Subfertility and child neurodevelopmental outcome: a socio-economical point of view
Subfertile couples who may or may not seek for ART or already undergo ART are, on average, older, higher educated and more often nulliparous compared to the general obstetric population. Consequently, these patients carry additional age-related and parity-related risks or some unknown factor intrinsic risk factors potentially affecting the offspring, in an indirect or direct way, that cannot automatically be attributed to the ART treatment itself, as was demonstrated in Chapter 6. On the other hand, children born to
subfertile couples, with or without the application of ART, may generally benefit from an advantageous socioeconomic status. The relative favourable socioeconomic position (~40% highly educated subfertile couples) that was observed in this thesis as well may be an explanation for the number of relatively high IQ scores of the studied children.

Subfertility and child neurodevelopmental outcome: subfertility causes and obstetrical events

Subfertility can be caused by a disturbance in any of several reproductive processes. It is likely that subfertile women who conceived without ART may have different fertility problems than women with fertility issues who conceive with ART. Additionally, it is expectable that each subfertility cause has its own potential consequences for child outcome, based on its origin. It is known that subfertility is associated with obstetrical problems such as pregnancy-induced hypertension or preeclampsia, placenta praevia, and increased perinatal adversities such as preterm birth, low birthweight and even perinatal deaths. The occurrence of perinatal deaths and prematurity indicate that subfertile women often have a poor overall reproductive function, as they had difficulties with both to conceive and to carry the pregnancy to term. On the other hand, one should realize that women with a history of subfertility are perhaps more closely monitored than others when finally a pregnancy is achieved, with or without the application of ART. Consequently, they have a higher pick-up rate of any potential problems, leading to an increased diagnosis of obstetric complications and the experience of stress, prior to and during the pregnancy.

In the present thesis, a relation between child developmental outcome and the underlying subfertility cause, in terms of tuba pathology, endometriosis, cervical factor or hormonal cause, male subfertility or unknown cause of subfertility could not be demonstrated. The performed subgroup analyses, however, reduced the sample size of our groups and may have affected the precision of the results. It may also be possible that other conditions than measured underlie the subfertility of the studied couples and interfered with child neurodevelopment.

Subfertility and adverse child neurodevelopmental outcome: parental and prenatal stress

Another interesting thought in the explanation for the interaction of subfertility with child development is its relation to parental and perinatal stress. During fetal life, the nervous system rapidly develops, making the fetal nervous system more vulnerable to environmental influences. As discussed in Chapter 5, the data may imply that suffering from prolonged subfertility triggers the neurodevelopmental vulnerability for both sexes, and with that, nullifies the sex-specific neurodevelopmental vulnerability of the boys. An explanation for the adverse effect of TTP, which was particularly seen in girls, may be that it is due to maternal stress to which the fetus is prenatally exposed. It is likely that couples experience subfertility as a psychological burden during the period of unwanted childlessness as well during the finally achieved pregnancy. However, couples that eventually achieve pregnancy after a long period of subfertility were able to deal with this long period of unwilling childlessness and many ART treatments, are presumably more
capable of coping with stress. Within the Groningen ART cohort study, it was previously demonstrated that a longer TTP was associated with less trait anxiety of 1-year-old children.\textsuperscript{201,242} Yet, the latter does not preclude that the offspring is affected by stress in utero as subfertile couples may underestimate the incurred stress as the joyfulness of having a child at last may overshadow the rough period of prolonged subfertility.\textsuperscript{202,203} Evidence is accumulating that maternal stress during pregnancy is associated with an increased risk of disturbance in offspring neurodevelopment.\textsuperscript{204,205} Cortisol and testosterone have been proposed as mediating hormones between maternal mental status and fetal development, as cortisol and testosterone are both the end products of two hormonal axes, and perinatal exposure to high levels of cortisol and testosterone is associated with perinatal adversities and health problems and may influence the stress response of the child itself. The altered neurochemistry due to exposure to prenatal stress may result in structural changes of the developing young brain. It is well-known that gonadal steroids play a crucial role in the sexual differentiation of the brain and its development. Animal studies have shown that prenatal stress masculinizes females and feminizes males, under the influence of androgens, impairing the reproductive capacity and reducing the sex-specific behaviour of the organism.\textsuperscript{206-210} Similar associations were seen in humans by Barrett et al. who found that exposure to prenatal life events stress was significantly associated with a longer, more masculinized anogenital distance (AGD, an indicator of prenatal androgen exposure) and masculinized play behaviour in girls, as well as a trend towards shorter, less masculine AGD in male infants but not with feminized play behaviour in boys.\textsuperscript{211,212} The authors suggest that prenatal stress may act as a non-chemical endocrine disruptor, interfering with sex-typical reproductive and neurodevelopment.\textsuperscript{211,212} The results of these studies support the hypothesis that the hormonal perturbations due to the parental subfertility-related stress mediate the neurological findings of the follow-up of the Groningen ART cohort children at 4 years: the hormonal changes may masculinize the female brain and induce more boy-like neurological vulnerability in girls, especially in the case of a long TTP, were parental stress is prolonged.

**Clinical implications for subfertile couples and their children**

*Counselling: being subfertile or suffering from subfertility?*

Making the transition to daily practice, what are the clinical implications of the results of this thesis? How to counsel couples with an unfulfilled child wish who may have a request for ART? Subfertile couples should be generally acknowledged as a potentially high risk group when providing prenatal care. Ideally, but maybe not the most realistic, (severe) subfertility should be prevented, which can be achieved to some extent by encouraging couples not to postpone having children, at least not until their mid-thirties. At any antenatal booking, a history of subfertility should be sought out. Counselling before any fertility treatment should include information on the potential consequences for the mother and the developing child. It should be mentioned that some of these potential consequences are inherent to the obstetric adversities itself, whereas others may be more directly related to
ART or the underlying history of subfertility. The decisions about the place of delivery for subfertile women may be influenced by this information.

Likewise, professionals must ensure that they do not unnecessarily unease the future parents, as the absolute overall risk of obstetric and neonatal complications is relatively small. Moreover, it is a challenge for the clinician to properly estimate today's effects of ART and subfertility as ART treatment protocols have changed over the years and continue to change.

The impact of minor neurological signs

The current literature available, including the current thesis, generally suggest that if ART or subfertility are related to neurodevelopmental outcome, the effect is expressed in the more subtle parameters of child neurodevelopment. Owing to their relative high incidence in a population, minor neurological signs, such as neurological non-optimality and MND, may serve as indicators for suboptimal development. In fact, the more subtle expression of the neurological condition may serve as a paradigm to understand the underlying pathophysiology in brain development related to internal or external factors. A more clinical implication of minor neurological signs is that of the presence of MND. The presence of MND, especially complex MND, is an indication for the child's vulnerability for developing co-occurring problems, such as specific learning disorders or attention problems. It has been demonstrated that MND is associated with impaired motor performance at school age.\(^{156,158,249}\) MND is also associated with learning disabilities, including spelling, reading and writing impairments and impaired arithmetic skills and in a lesser extent to behavioural problems, such as attention problems, social problems and internalizing and externalizing behaviour.\(^{90,157,199,200,218,250-252}\) The findings of this thesis also suggests the presence of interrelationships between MND and different types of developmental disorders. A clear association was found between increased TTP and both the occurrence of complex MND and lower IQ scores at 4 years and between the presence of subfertility and behavioural problems. Additionally, cognition and behaviour were inter-related. The complex variant of MND is of clinical relevance due to its clear association with learning and behavioural problems.\(^{90,157}\) Moreover, from an aetiological perspective, complex MND can be considered as a borderline variant of cerebral palsy (CP), given a corresponding association between complex MND, CP and perinatal adversities.\(^{90,253}\) Simple MND can be considered as a mild variant of MND, reflecting a typical yet non-optimal development of the brain. Simple MND is frequently found in children and may be regarded as a minor neurological difference rather than a dysfunction. Unlike complex MND, simple MND is not evidently related to adverse conditions during early life.\(^{90}\) Although simple MND is considered to have little clinical relevance, it is notable that simple MND, like complex MND, is associated with an increased risk for cognitive impairments and behavioural problems in school-aged children.\(^{199,200,252}\) In addition to the clinical neurological condition in terms of MND, the derived optimality parameters, the fluency score and NOS, are included in the studies of this thesis. The optimality scores quantify subtle neurological deviations in the neurological
condition of a child whereas the classification of MND is based on the presence of coherent clusters of neurological deviant signs. Clinical neurological condition and neurological optimality are highly correlated; however, a reduction in neurological optimality does not necessarily mean the presence of a neurological deviation, as the range for optimality is narrower than that for normality. The quantitative nature of the optimality scores makes it a suitable instrument to evaluate subtle neurological deviations for scientific research purposes rather than for clinical purposes.

As described above, minor neurological signs often do not have clinical relevance for an individual, but, on the level of a population, they may be of significance. Currently, up to 5% of newborns in Europe and 1% of newborns in the United States are born following ART. In the Netherlands, up to 3% of children is born following ART. Theoretically, this means that in Europe, at least one child in every classroom has been born following ART and even more children may be the result of natural conception of subfertile couples. Wittingly or unwittingly, everyone is familiar with a former classmate, who was always chosen last during gym classes. This classmate was moderately popular on school or even the misfit of the class and turned out to be a repeater. Other children made fun of him or her because of his or her clumsiness or sometimes peculiar behaviour. You may have watched these events from a distance, stood up for this person or this person was actually you. Whatever the exact situation was, it is conceivable that this particular child had complex MND, given the clumsiness in combination with the behavioural problems and the poor school performance. It goes without saying that the events in the case have an impact on the self-esteem and psychosocial functioning of young children that may persist during puberty and thereafter. The case also illustrates that subtle neurological dysfunctions may interfere with academic achievement and activities in daily life.

**PGS and child neurodevelopmental outcome**

This thesis concludes that PGS does not seem to affect neurodevelopmental outcome of 4-year-old singletons. In contrast, PGS does seem to affect neurodevelopmental outcome of 4-year-old twins: PGS in twins was associated with a negative effect on neuromotor condition and a positive effect on the cognitive function sequential processing. This may point to the possibility that the embryo biopsy inherent to PGS is associated with differences in brain function at a later age. The findings correspond to some extent to previous findings of the PGS study. At the age of 2 years, singletons and twins were analysed together, which means that a potential effect of PGS could not be attributed exclusively to singletons or twins. It was demonstrated that children born after IVF with PGS had, on average, an approximately two points lower NOS than controls. At 4 years of age a similar effect was found in PGS twins, who scored approximately two points lower on the fluency score and approximately four points lower on the NOS compared to the controls. However, such an effect was not present in singletons. A negative effect of PGS was observed for singletons by Seggers et al., who investigated the same group of 4-year-old children as part of the PGS study. The authors found a higher frequency of received
paramedical care (speech, physical or occupational therapy) in both PGS singleton and twins compared to controls.

Within the current literature available, the PGS study is the only research project in which the effect of PGS on child outcome is studied. Other studies in the field did not differentiate between PGD and PGS, while the indication to perform one or the other is rather different. Couples who are considered for PGD treatment are at a high risk of having a child with a genetic disorder, but do not have fertility problems, whereas PGS couples are not at risk for a genetic disorder, but do have fertility problems. Hence, it is challenging to compare the findings of the PGS study to other studies. With regard to singleton children, two other groups reported on developmental outcome of singletons born after embryo biopsy. Nekkebroeck et al. reported on child mental, motor, socio-emotional and language development related to PGS/PGD at the age of 2. The authors demonstrated similar outcome in singletons born following PGD/PGS and singletons born following ICSI or natural conception. Banjeree et al. reported on health and developmental outcome, measured with the Griffiths scale, in singleton children born following PGD/PGS up to the age of 4 years. Health and developmental outcome of the PGD/PGS children was similar to that of naturally conceived peers, with the exception of locomotor development, where PGD/PGS children scored significantly lower than the controls.

The current PGS study reported that PGS affected neurodevelopmental outcome of twins. This, however, does not match other studies. Nekkebroeck et al. were not able to demonstrate differences in mental, motor, socio-emotional and language development between 2-year-old PGS/PGD twins and twins born following ICSI or following natural conceptions. The authors demonstrated similar outcome in singletons born following PGD/PGS and singletons born following ICSI or natural conception. However, the results from the current thesis match to some extent to those of the study of Liebaers et al., who found higher rates of prematurity and low birthweight in PGD/PGS multiples than in ICSI multiples and more perinatal deaths in post PGD/PGS multiple pregnancies than in post ICSI multiple pregnancies, whereas outcomes were similar for PGD/PGS and ICSI singletons. Yet, in a larger analysis by Desmyttere et al., including the analyses reported previously by Liebaers et al., no differences in prematurity, low birthweight and perinatal deaths were found between PGD/PGS multiples in comparison with ICSI multiples. This means that the effects of PGD/PGS on child outcome are inconsistent. Also our results do not necessarily suggest an overall adverse effect of PGS on outcome of twins, as PGS was associated with a negative effect on neuromotor condition and a positive one on sequential processing, an indicator of short-term memory. Our results rather suggest that embryo biopsy inherent to PGS affects neurodevelopmental outcome of twins in a different way than that of singletons. Generally, twins are at increased risk for developmental problems. Twins may be more vulnerable to external factors that may influence brain development. Moreover, they may be more subjected to such adverse circumstances. For instance, twins share the uterus, which may not be the ideal intra-uterine environment for embryonic and fetal development.
Several explanations may be offered for a potential interference of PGS with the young, developing human brain. First, the removal of one or two blastomeres may negate any benefits that otherwise might derive from screening as it may comprise the developmental potential of the embryo with potential long-lasting consequences for the developing human. Second, the embryo biopsy inherent to PGS may not be a proper instrument to select normal embryos with a good developmental potential. One of the limitations of the technique is the limited number of chromosomes that can be analysed with FISH. Consequently, an embryo that was initially labelled as normal could be aneuploid for one or more chromosomes that were not tested. Another issue is that in assisted conception, chromosomal abnormalities in human embryos seem to be the rule rather than the exception even in morphologically good embryos and irrespective of maternal age. This phenomenon is most likely due to mosaicism – i.e. the occurrence of more than one cell line with different genotypes in one individual. As a consequence, there is a reasonable chance that the chromosomal constitution revealed by analysis of the blastomere(s) may not be representative for the entire embryo. In fact, PGS may even favour the selection of an abnormal embryo that screens normal, over the morphologically superior embryo that screens abnormal. One factor underlying the high rate of aneuploidy and mosaicism may be the ovarian hyperstimulation. In a trial among low-risk couples that compared mild ovarian stimulation and conventional ovarian hyperstimulation, the percentage of abnormal embryos relative to the number of embryos diagnosed was 45% following mild stimulation (40 patients) compared to 63% following conventional stimulation (40 versus 33 patients; \( P = 0.02 \)). In order to gain insight into chromosomal mosaicism, the authors then analysed the group of embryos in which two cells were available for diagnosis. Overall abnormality rates (abnormal and mosaic embryos) were 55% following mild and 73% following conventional ovarian stimulation (38 versus 30 patients; \( P = 0.046 \)), confirming the difference in abnormality rates observed after single-cell diagnosis. Interestingly, the proportion of mosaic embryos per patient was more significantly increased following conventional ovarian stimulation (65 versus 37%; \( P = 0.004 \)). In the present thesis, the application of ovarian hyperstimulation in both the PGS study group and the control group may be one of the explanatory factors for the similarity in neurodevelopmental outcome of the singletons.

The results described in Chapter 7 are not directly alarming, however, it may be possible that PGS, a procedure involving embryo biopsy, has any late consequences for brain function at a later age. Even though PGS is no longer practised on routine basis, the need for careful monitoring of children born following embryo biopsy remains. Embryo biopsy is still performed in the form of PGD and new techniques are continued to develop.
Methodological considerations
Scientific findings must be considered within the context of their methodological limitations. Ideally, a double-blinded randomized clinical trial would be performed on the effect of assisted conception and subfertility on child outcome. However, randomly allocate (sub)fertile couples to assisted conception or natural conception is not ethical.\textsuperscript{54,262} The second best study design to thoroughly explore the effect of assisted conception and subfertility may be in the form of a control study. Difficulties to overcome in such design are namely the extended period of time it takes to conduct a long-term prospective cohort, the number of participants lost to follow up, the choice of a control group and the definition of the evaluation criteria. The methodology of the two parallel running projects of this thesis, Groningen ART cohort study and the PGS study will be addressed in the following paragraphs.

The Groningen ART cohort study
The major strength of the Groningen ART cohort study lies in its unique design. First, the composition of the study groups made it possible to study the effects of separate components of assisted conception related aspects on the child’s neurodevelopmental outcome. Children born to couples who underwent ovarian hyperstimulation IVF (COH-IVF) were compared to children born to couples who underwent IVF in a modified natural cycle (MNC-IVF), without ovarian hyperstimulation. In turn, both groups were compared to a control group, a group of children born to subfertile couples who underwent no fertility treatment but conceived naturally. Parental characteristics of the latter group more closely resembled those of children born to COH-IVF and MNC-IVF, rather than a fertile control group of the general population had. Such comparison reflects the effect of ovarian hyperstimulation and the in vitro procedure together and is confounded by parental characteristics and factors associated with subfertility. As a result of the composition of the Groningen ART cohort study groups, the effects of potential confounders and a potential overestimation of the effect of ART were minimized. Second, the prospective design of our study reduced potential selection bias based on the child’s development or health, as we invited the couples of the Groningen ART cohort in the third trimester of pregnancy. Third, the assessors and supervisor were blind to prenatal and perinatal history, including the mode of conception of the ART study groups. Moreover, the likelihood that the assessors or supervisor would guess the conception mode was minimized, because all ART cohort couples had experienced subfertility. The effort that was put in the maintenance of couples with their children in the follow-up resulted in a representative study sample up until the assessment round at age 4. Throughout the study years, the postnatal attrition rate has reached only 9.3%, with an initial prenatal enrolment of 63% to 76% of eligible children.

Standardized, age-specific and highly sensitive measurements were applied to evaluate neurodevelopmental outcome in children ranged from 4 months to 4 years of age. With the Hempel examination and the IMP not only traditional signs of neurological dysfunction such as muscle tone dysregulation and motor delayed milestone achievement are assessed, but
the quality of motor behaviour is taken into account as well. Subtle changes in neurological development are relatively unimportant on an individual level; however, on a population level they are relevant, since ART children represent a substantial part of society. Most studies however use relatively gross measurements on neurological outcome, such as the Bayley’s or Griffiths Scale of Infant Development. When instruments are not designed to study neurodevelopmental outcome in a detailed sense, potentially subtle differences between groups may remained undetected. For instance, Bouwstra et al. demonstrated an adverse effect of neonatal trans-fatty acid status on neurological development with the use of the detailed and sensitive Hempel assessment but not with the use of the Bayley’s Scale of Infant Development.164

In all studies in this thesis, several statistical approaches were applied in order to approximate the truth as closely as possible. In Chapter 2 and 7, mixed models were used as statistical model which was helpful analysing the repeated measurements on the IMP at 3 different ages and helpful analysing the child-to-mother related neurodevelopmental measurements in the PGS study. In Chapter 3, 4 and 5 univariable and multivariable regression analyses were applied to estimate the relationship between ART-related factors and subfertility aspects and neurodevelopmental outcome. In the multivariable analyses adjustments were made for potential confounders on the pathway from assisted conception, subfertility and neurodevelopmental outcome, such as birthweight, gestational age, sex, vanishing twins and parental age and educational level. The consideration to adjust for these potential confounders was twofold: a priori, based on the available literature and based on significant differences in background variables between groups. In Chapter 6 causal inference search algorithms and structural equation modelling were applied as statistical tools. Unlike traditional statistics these methods are able to unravel underlying causal mechanisms and distinguish between confounders and intermediate effects.

No study goes without criticism, therefore some points need to be discussed regarding the Groningen ART cohort study. One limitation of the study is related to its power. The original power calculation was based on the neurological outcome at 18 months.140 Moreover, the calculation was based on the neurological outcome in the ART study groups and not based on possible associations between TTP and neurological outcome. Lastly, the neurological outcome parameters to calculate the study size were based on the Hempel examination and not based on the Infant Motor Profile (IMP). At baseline, the size of the MNC-IVF group was somewhat smaller than the 64 needed for an adequate power of the study. Partly, this was compensated by larger groups of COH-IVF and Sub-NC children and fortunately, the sample size of the study groups did hardly decrease throughout the study years. In order to deal with the variety in outcome parameters and the association between TTP and neurological outcome two post hoc power analyses were performed. These analyses demonstrated that the sample size of the ART groups were slightly too small to detect ART-related effects on the IMP variation score. With respect to the TTP-effects on the prevalence of complex MND at 4 years, the sample size of the subgroup of girls was
large enough to detect relevant effects, whereas the size of the total group was somewhat underpowered.

Another limitation is that the mediation used in MNC-IVF treatments, although minimal and no hyperstimulation medication, may have caused an overestimation of the effect of the in vitro procedure or an underestimation of the effect of ovarian hyperstimulation. However, this minor confounding effect of MNC-IVF did not complicated the interpretation of the study results, as no neurodevelopmental differences between the three ART study groups were found. Another aspect of the MNC-IVF group should be mentioned. Women who underwent MNC-IVF fulfilled specific selection criteria with regards to age, BMI, previous ART treatment and the presence of an ovulatory menstrual cycle. Although we have put effort in reducing potential selection bias, it is conceivable that subtle group characteristics other than the aforementioned could have had interfere with the actual effects of assisted conception and subfertility.

Another limitation is the composition of the fertile reference group at the 2 and 4 years follow-up round. Recruitment of the reference group does not match the way in which we recruited the children of the Groningen ART cohort. Parents volunteered for participation in the reference group, which could have introduced selection bias. Parents with concerns about the health and development of their child are presumably more willing to participate in a developmental study. Moreover, the assessors and supervisor were not blinded to the group status of the reference children since this group was newly recruited.

It may be considered as a limitation that no distinction was made between IVF with ICSI and IVF without ICSI. The primary aim of the Groningen ART cohort study was to unravel the effects of ovarian hyperstimulation and the in vitro procedure on neurodevelopmental outcome and not the effect of ICSI in addition to IVF. Subdividing the ART study groups in ICSI-offspring and non ICSI-offspring would have reduced the sample size and subsequently, decreases the likelihood of finding certain effects. Once, however, attention was paid to the effect of ICSI, in an analysis in which the COH and MNC groups were pooled. This analysis did not reveal an effect of ICSI on the optimality scores and the prevalence of complex MND in 4-year-old singletons (Chapter 5).

It may also be considered as a limitation that only singletons were studied. Consequently, the results cannot be generalized to multiples, as a substantial part of those are born following controlled ovarian hyperstimulation IVF. Multiple gestation is associated with perinatal adversities and multiples are more prone to develop developmental disabilities. It is therefore legitimate to study developmental outcome for singletons and twins separately.

**The PGS study**
The PGS study is the first follow-up study of a randomized controlled trial on PGS in which children were followed as long as 4 years. The random assignment of parents resulted in a strong resemblance of the two groups with respect to most background variables so that comparisons between both groups primarily reflect the effect of PGS. Additional strengths
with regards to study design are the blinding of the assessors and supervisor to the mode of conception and the prospective design of the study. Another strength of the present study is the examination of both singletons and twins, which contributes to the generalizability of the study. Like the Groningen ART cohort study, standardized, sensitive and age-specific measurements were applied to evaluate neurodevelopmental outcome.

In both PGS and control groups children born after IVF with and without ICSI were included. In contrast to PGD where only IVF with ICSI is applied, IVF with or without ICSI are both applied whenever PGS is indicated or not. However, it could be argued that ICSI substantially differs from IVF alone. The minimization for reproductive technique (IVF or ICSI) during enrolment and the application of a sensitivity analysis dealt with the problem, indicating that the presence or absence of ICSI did not modify a potential effect of PGS.

The major limitation of the PGS study is the relatively small sample size. Power calculation was based on the number of women needed to detect a certain increase in the cumulative ongoing pregnancy rate. The unforeseen effect of lower pregnancy rates after PGS further reduced the number of children available for follow-up. A post hoc sample size calculation indicated that the study groups were adequately powered for the majority of outcome parameters.

Another limitation is the selective drop-out in both groups (drop-outs PGS group: higher gestational age; control group: lower educated parents). The selective loss of lower educated parents in the control group may not adequately explain the different neurodevelopmental outcome of PGS twins compared to control twins, as the selective loss was present in twins and singletons, where the selective loss in the singletons was not associated with developmental differences between PGS children and controls. Moreover, an explorative analysis revealed among others no statistically significant effects of parental educational level or gestational age.

It should be noted that participating children of both groups turned out to have a relatively high IQ. An explanation for this finding may be the relatively large proportion of highly educated parents participating in our study, as it is well known that cognitive abilities of children are positively associated with their parents education level. This means that our findings have to be interpreted with caution.

**Future perspectives**

The general aim of this thesis was to evaluate the effects of assisted conception and subfertility-related aspects on neurodevelopmental outcome in preschool-aged children. The body of evidence on long-term health and development of children born to subfertile couples, with or without ART, derived from good methodological quality has to grow further. The results described in this thesis need confirmation and neurodevelopmental follow-up has to be extended to school age and the phase of adolescence to elucidate the long-term effects of ART and subfertility on the developing brain against the light of puberty-onset. A study that may confirm the findings of the current thesis, is the INeS study. Some years ago, this randomized clinical trial was conducted to investigate
whether COH-IVF and MNC-IVF can reduce the number of multiple pregnancy rates, but uphold effectiveness, in terms of a healthy singleton, costs and patient preferences in comparison to intra-uterine insemination (IUI) with ovarian hyperstimulation (COH-IUI). Couples with unexplained or mild male subfertility were allocated to either 6 consecutive cycles of COH-IUI or 6 cycles of MNC-IVF or 3 cycles of COH-IVF with single embryo transfer (eSET) plus subsequent cryo-cycles. Although beyond the initial aim of the INeS study, the composition of the study groups and the randomized character of the study offers the possibility to thoroughly study the effects of ovarian hyperstimulation, the in vitro laboratory procedures and the combination of both on obstetric and child health and developmental outcome, with minimization of selection bias. The potential effect of ovarian hyperstimulation can be further explored by comparing embryonic and endometrial development after fresh and frozen-thawed embryo transfer.

Next to neurodevelopmental outcome, health outcome parameters such as cardiometabolic outcome, respiratory health and cancer deserve further attention, as recent evidence has emerged that ART children have a distinct metabolic profile that may predispose them to cardiovascular disease later in life.

**Upcoming reproductive medicine: single embryo transfer and frozen embryo replacement**
An important step ahead in reproductive medicine is to increase the application of eSET and with that decreasing the risk of achieving a multiple pregnancy and its associated adverse effect for the mother and the child. In eSET, a single high-quality embryo at either the cleavage or blastocyst stage of embryo development is selected from a larger number of available (high quality) embryos for transfer to the uterus. Currently, the number of multiple embryo transfers and multiple delivery rates are declining, while fortunately, the reduction in the number of transferred embryos has not lowered the overall pregnancy rates. At present, approximately 25% of embryo transfers in Europe result from eSET, whereas approximately 50% of transfers result from double embryo transfer (DET) and in less than 20% of cases, 3 or more embryos are simultaneously transferred to the uterus. The current trend of transferring fewer embryos has resulted in more embryos being available for freezing (frozen embryo replacement, FER). The increased use of FER has intensified the awareness of the safety aspects of this procedure and also offers the opportunity to gain more insight in the effect of the supraphysiologic hormonal milieu resulting from ovarian stimulation. In contrast to fresh embryos, cryopreservation embryos are not exposed to a hormonally loaded endometrium, presumably resulting in more optimal endometrial circumstances. Support for this hypothesis is promoted by Wennerholm et al., who systematically reviewed the literature available on child outcome after cleavage stage embryo cryopreservation up until 2008. The authors concluded that children born after FER showed similar or even improved perinatal outcomes compared to children born following fresh cycles. Moreover, the proportion of low birthweight and prematurity after FER is similar to that of singletons born after from spontaneous conception. After the review of Wennerholm et al., two Scandinavian large cohort studies
were performed on perinatal outcome of children born following FER in comparison to peers born following fresh embryo transfer. Like Wennerholm et al., the studies concluded that FER does not adversely affect perinatal outcome and the outcomes are similar or even better, particularly regarding fetal growth, compared to those of children born following fresh embryo transfer.\cite{176,269} However, there are concerns about an increased risk of being large for gestational age (LGA) after FER in comparison to both fresh embryo transfers and spontaneous conception.\cite{269,271} It is hypothesized that the cryopreservation techniques may induce changes in the developmental processes in the early embryo stages and hence in the intra-uterine growth potential. The possible asynchrony between the endometrium and embryo that occurs in case of FER may also play a role in the occurrence of higher rates of LGA children after FER as it may alter the subsequent trajectory of fetal growth and development.\cite{33,271} The increased risk of being LGA after birth following FER deserves further attention, given the growing trend towards single embryo-transfer policy and the resulting increase in freezing of the non-transferred embryos.

The future of preimplantation genetic screening

In modern society, science is driven at top speed and also in reproductive medicine new developments in assisted conception take place. Such fertility treatment changes may have an impact on the various outcomes and complicate the interpretation of study findings. It appears that over time, the overall outcomes for children born following ART has improved.\cite{272} The improvement may be the result of the enhancement of clinical and laboratory skills and the increasing use of eSET or the current accessibility of ART for subfertile couples. With the introduction of new treatment regimes, responsibility follows to carefully monitor and warrant its safety. Ideally, safety evaluation is performed prospectively. However, from an ethical point of view, such strategy is usually not possible in reproductive medicine. Consequently, theoretically well-thought techniques may turn out differently in daily practice, as in the case of PGS.\cite{22,23} PGS as applied in the PGS study in this thesis is currently no longer practiced, however, embryo biopsy is still applied in the form of PGD. Moreover, the techniques used in PGS and PGD are subjected to developments to improve its accuracy in terms of the timing of the biopsy (i.e. cleavage or blastocyst stage), the material used for screening (embryonic cells, the polar body or trophectoderm tissue) and the number of chromosomes that are needed to be screened. At present, it is possible to screen for aneuploidies using all 23 pairs of chromosomes of a single cell using more advanced genetic analysis such as comparative genomic hybridization arrays or single nucleotide polymorphism arrays.\cite{273} Future research should focus on the effectiveness of these new techniques. And potential adverse effects on health and development of the offspring has to be identified before large scale implementation of new techniques.
Extending long-term follow-up: are subfertile couples becoming grandparents?

It is important to investigate whether adverse health and developmental outcome of children manifest later in life. It is to be expected that in the coming years, the follow-up period of children born to subfertile couples who conceived spontaneously or with the help of ART will be expanded to adolescence and adulthood as an increasing number of this next generation has reached the fertile age. The first person born following ART, Louise Brown, is now 36 years old. She had her first child at the age of 27 years, a son named Cameron, who was conceived naturally within 6 months TTP.\textsuperscript{274,275} Whether or to what extent some subfertility problems are heritable may depend on the type of subfertility. For instance, Louise Brown’s mother underwent IVF because of blocked fallopian tubes, which is usually an acquired condition and may therefore not influence the fertility patterns of the offspring.\textsuperscript{6,9} However, other fertility outcomes may be expected in case of unknown subfertility, male subfertility or subfertility due to polycystic ovarian syndrome or endometriosis, disorders that may already be inheritable by itself. Boys conceived using ICSI inherit the Y-chromosome of their fathers, which may contain abnormalities that may be related to the underlying male subfertility of the father. Whether ART interferes with the fertility potential of the offspring is an interesting question that needs further research. In theory, ART may interfere with the early development of the endocrine organs due to hormonal stimulation, which may be reflected by disturbances in the pubertal development. From the scarce literature available exploring this topic so far, it can be deduced that ART does not appear to affect the onset of puberty.\textsuperscript{75,85,86} However, there are some suggestions that ART may interfere at a different level in puberty, such as delayed breast development, advanced bone age and higher serum levels of luteinizing hormone (LH) and dehydroepiandrosterone sulphate (DHEAS, a steroid hormone, that functions as a metabolic intermediate in the biosynthesis of the androgen and oestrogen)\textsuperscript{85} in girls and lower serum testosterone concentrations and higher LH/testosterone ratio in boys.\textsuperscript{88} In the latter study, the deviant hormonal levels were only seen in boys born following ICSI, but not in boys born following IVF or natural conception, suggesting that the ICSI boys may have inherited the subtle impairment of Leydig cells from their fathers.\textsuperscript{88} However, another study reported normal male prepubertal development after ICSI, in terms of penile and testicular growth as well as Sertoli cell function.\textsuperscript{276} During puberty, the same group of ICSI boys showed testosterone levels comparable to NC boys.\textsuperscript{277,278} Moreover, ICSI boys from fathers with severely compromised spermatogenesis showed testosterone levels comparable to those from fathers with normal spermatogenesis.\textsuperscript{277,278}
Concluding remarks

The following conclusions can be drawn from the studies described in this thesis:

- Ovarian hyperstimulation, the in vitro laboratory procedures, or a combination of both are not associated with worse neurodevelopmental outcome - in terms of movement variation, neuromotor development, cognitive and behaviour development - in singleton children aged 4 months to 4 years.

- The severity of subfertility, reflected by an increased TTP, rather than the presence or specific underlying causes of subfertility are associated with minor neurological dysfunction at 2 and 4 years.

- The presence and especially the severity of subfertility affected cognitive and behavioural development at 4 years, with maternal age at child conception and maternal educational level as the most important confounders.

- Neurodevelopmental outcome – in terms of neuromotor development, cognitive and behaviour development – in singletons born following IVF with PGS is similar to that of singletons born after IVF without PGS at age 4. However, PGS seems to be associated with altered neurodevelopment in twins.

The main conclusion that subfertility-related aspects rather than ART-related aspects contribute to a worse neurodevelopmental outcome of preschool-aged children, with the severity of subfertility being the decisive factor, underlines the importance of long-term follow-up of health and development of children born to subfertile couples, without and with the application of ART. In a society where maternal age at child birth, subfertility and the application of ART are steadily increasing, it is of utmost importance to monitor and to improve the safety and health of subfertile couples in their quest for parenthood and of their future offspring.